RECOMMENDATIONS FOR DEPLOYING A PORTAL FOR NASA'S WORKFORCE AND THE PUBLIC

Jeanne Holm, Team Lead
Chris Pino
Douglas Hughes
Steve Prahst
Justin Jackson
Jayne Dutra
Brian Dunbar
Andy Schain
JoAnne Rocker
Lisa Nayman
Robin Land
Steve Naus
Chan Kim
Manjula Ambur
Vicki Pendergrass
Nancy Kaplan
Maria Chacon
Bill Price

November 28, 2001

National Aeronautics and Space Administration

EXECUTIVE SUMMARY

NASA's information resources are vast—more than 2,000,000 web pages, thousands of databases and electronic repositories, petabytes of mission and planetary data, and millions of online reports—a wealth of electronic information. Unfortunately, trying to navigate, access, and sort all that information is often time consuming, difficult, and discourages people from seeking out or sharing lessons learned. In addition, our web presence is our most often accessed image to the public. With close to 2 *billion* hits a month on NASA sites, the web far surpasses other methods of interacting with the public and with our internal workforce.

One of the primary ways in which to both manage the information on an organization's web space and to provide better, faster access to that information is through the use of a "portal". A portal is an electronic gateway that offers easy access to online resources through a personalized home page that collects links, headlines, and business information most relevant to the user.

While the portal can deliver a clear return on investment within a model of helping to manage NASA's web space, there are additional goals that the portal will help to achieve for NASA's communities as noted below.

- **Public**
  - Support an integrated view for the public into the diverse face of NASA
  - Organize and manage the NASA electronic resources to deliver the intellectual power of NASA to educators, families, and citizens
- **Employees**
  - Increase productivity by facilitating quick access to and sharing of information across organizational and discipline boundaries
  - Create a broader sense of community through sharing news and successes
- **Teams**
  - Allow virtual teams to quickly share and learn from others, while building a legacy for future projects and programs by capturing key team decisions

- **Partners**
  - Support management of the end-to-end information flow with our suppliers and outside partners
  - Securely and appropriately share the most practicable information with our partners to ensure clear communications and better decisions

Studies have shown that if someone has to search more than 15 minutes for a piece of information, then they are highly likely to simply re-invent the item (a drawing, specification, or finding). Such re-invention encourages errors, increases mission risk, and wastes time that could be better spent in research or project support.

Content is much more than just a listing of links to static web sites. Content is what brings people back to a portal, over and over again. It's what causes them to bookmark it or make it their home page. The content of the portal needs to be informative, accurate, timely, and entertaining. Content management looks at identifying and streamlining our publication processes, understanding where our knowledge resources exist, and capturing and archiving electronic information as it is created and shared.

Portal-related activities have been underway for some time at NASA. Targeted portals for specific communities are currently in operations, for example the Technology Portal for Code R, MyGoddard and InsideJPL as Center-wide portals, and the Process-Based Mission Assurance Knowledge Management System for Code Q. The portal, when deployed, will allow information to flow more freely across the entire organization, bypassing archaic or inadvertent barriers that currently exist. Concepts of the look-and-feel of an InsideNASA portal are shown in Figure A.

Although white papers often stop short of making actual implementation recommendations, that is not the case here. The team members have extensive experience in many aspects of the web, and in bringing up operational portals at NASA. The recommendations at the end of this report address how to deliver the portal framework in the architecture shown in Figure B. Those components highlighted in blue (the Portal Framework box and functionality) are the focus of this white paper and implementation recommendations.

Figure A. InsideNASA Concepts

- **Phase 1 will deliver within 10 months to 2000 users**
    - The InsideNASA *and* MyNASA portals with 5-10 data channels each (only unrestricted data will be presented in Phase 1)
    - A first taxonomy and core metadata recommendations
    - The first content management processes for sheparding electronic information along its lifecycle, from creation to approval to archive

At the end of Phase 1, a review will be held to both document the lessons learned, incorporate them incorporate Phase 2 plans, and to assess a competitive procurement for a long-term portal and content management application. A serious consideration in this effort will be the maturity of the IFMP portal from SAP.

- **Phase 2 will add over the next 18 months Agency-wide capacity to provide**
  - Content management capabilities and a richer taxonomy
  - Increasing support to communities of practice through creation and facilitation of additional data channels
  - Security to allow integrated access to restricted and unrestricted information
  - Integration sign on to decrease the number of passwords an individual needs
  - Gradual scaling up to Agency-wide deployment for internal and external audiences to ~1,000,000 hits per week (based on analysis of current server logs). (During mission events, this number can reach 10,000,000 hits per hour.)
  - Based on experience with the JPL portal and expected fluctuations in demand for access by the public, we recommend that Phase 2 be hosted at a secure managed service provider.

The portal management will follow standards for software development, deployment, and operations. Over the past 18 months, the Jet Propulsion Laboratory (JPL) has brought up a portal. Partially funded by Code AO as a potential pilot for an Agency portal, this activity was very successful during prototyping. The portal management methodology is based on lessons learned by that team[1] and other portal development teams at NASA and in industry, as well as best practices.

This portal deployment is planned to be a portion of NASA's Web Management Services[2] and, as such, will be managed, funded, and conducted under that team's leadership. That team will need to work closely with other activities as shown in Figure B, the many distributed content providers, the SRRs related to secure nomadic access and publications policies, and teams such as the CIO community, Knowledge Management, and IT Architecture. In addition, coordination with other portal activities in development or operations is critical.

NASA is one of the world's premier institutions for knowledge creation and we should bring our knowledge dissemination services and techniques up to the same excellent level.

---

[1] Jayne Dutra, Sauwan Leung, Peter Scott, Douglas Hughes, and Charles Rhoades, *Inside Inside JPL: Project Review and Software Evaluation of the JPL Enterprise Information Portal Prototype*, Jet Propulsion Laboratory, Pasadena, July 27, 2001.

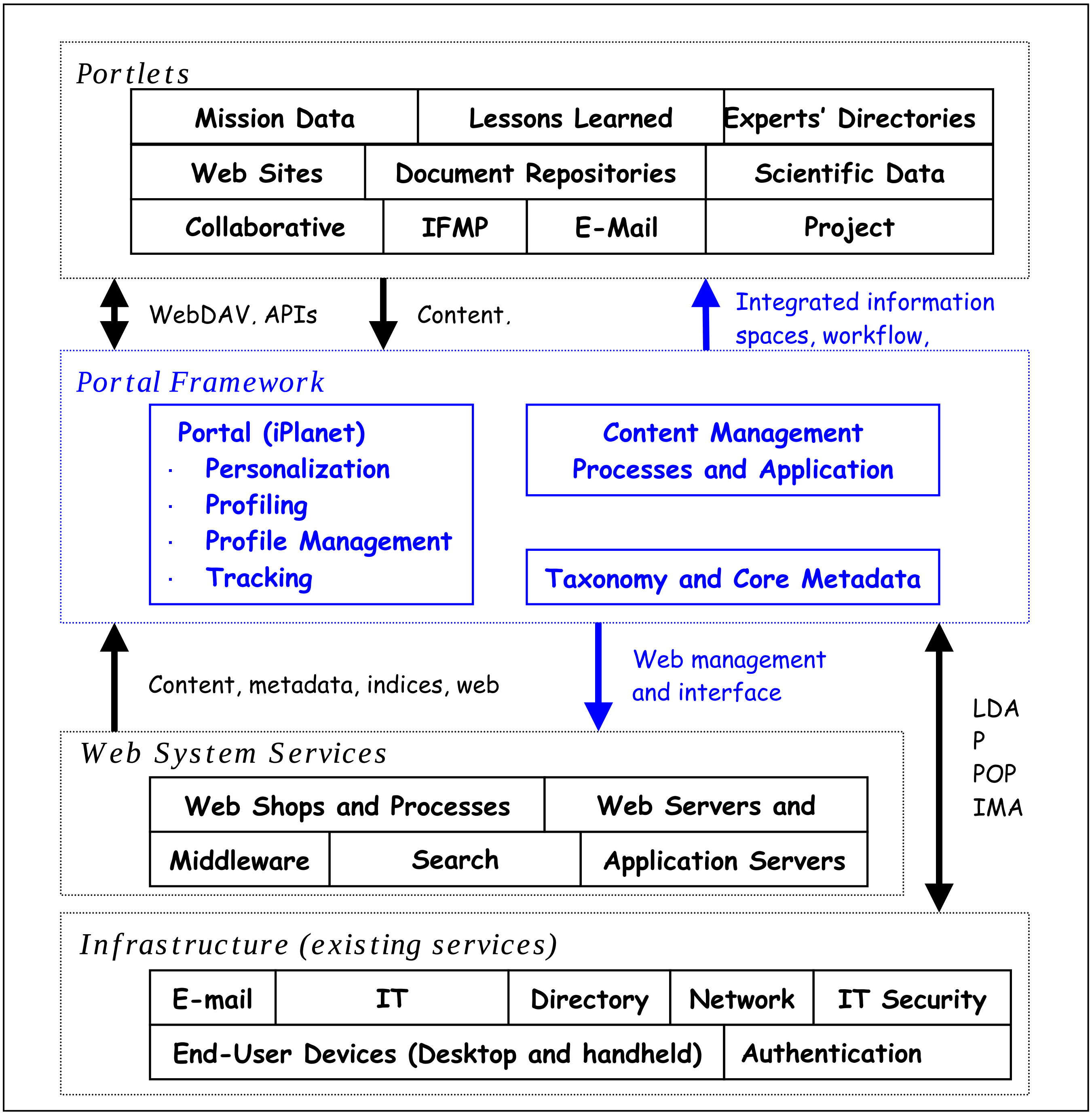

Figure B. NASA agency-wide portal architecture. This report focuses on those activities highlighted in blue (portal framework).

[2] The Web Management Services Team is led out of Code AO and was initiated as part of Strategic Resource Review (SRR) 67 in September 2001.

This report was created collaboratively by many talented individuals. Special recognition is due to the teams who worked on each section. Team leaders are shown in italics.

- **Requirements**: *Chris Pino*
- **System Architecture**: *Steve Prahst*, Douglas Hughes, Andy Schain, Brian Dunbar, Chan Kim, Chris Shenton, Bill Price
- **Information Architecture**: *Jayne Dutra*, Lisa Nayman, Andy Schain
- **Content Management and Business Processes**: *Justin Jackson*, JoAnne Rocker, Jayne Dutra, Lisa Nayman, Robin Land, Brian Dunbar, Nancy Kaplan
- **Portal Management**: *Douglas Hughes*, Steve Prahst
- **Implementation Plan**: *Jeanne Holm*, Douglas Hughes, Manjula Ambur, Jayne Dutra

TABLE OF CONTENTS

1. INTRODUCTION

Think of a portal as *your* view of interesting and relevant information. A portal is an electronic gateway that offers easy access to online resources through a personalized home page that collects the links, headlines, and business applications most relevant to you. Enterprise portals are intended to be one-stop information shopping to support communities of knowledge workers who share common goals.

A portal is a single, integrated point of comprehensive, ubiquitous, and useful access to information, applications, and people.

As we become increasingly reliant upon electronic resources, portal technology is going to become essential to NASA's ability to conduct its business. Why? Simply put, the ironic consequence of the explosion of available information is that it is becoming harder to find relevant information easily.

When NASA was smaller, and methods of publication were much more limited, it was relatively easy to collect, index and offer for distribution important information. A NASA employee could go to a technical library and be reasonably sure she was finding all information relevant to her query, usually in printed journals, books, or diagrams.

With the growth of the NASA organization and the advent of the Internet—with its ability to make every user a publisher—the search for information has become more complex. Who has the information? Has it been translated from hard copy into electronic format? In either format, has it been absorbed into a central collection? Is the information accessible from a user's desktop computer, or must a physical trip be made to access it? If the latter, is the information available at the user's Center in any form? Will online information require access to specific software? If so, has that software become obsolete?

These questions become even more complex considering the diversity of NASA's audiences. To use the model developed by the eNASA Team[3], NASA's audiences can be formed into four communities:

- Public (including general public, news media, students, and educators)
- Employees

[3] More information on eNASA can be found at http://enasa.ksc.nasa.gov

- Partners and suppliers (including contractors and international partners)
- Virtual teams (including contractors, civil servants, and academic institutions)

Each of these audiences is seeking different kinds of information. Even when pursuing information on the same topic, members of the difference audiences will want information in different formats (e.g., press releases versus scientific papers) or at different levels of detail (processed data sets versus factual summaries for use in schools).

As an added complication, an individual can easily become a member of more than one community, depending on the kind of information being sought. For example, a NASA civil servant may be an employee early in the day, so as to check online financial data about his project; the switch to Virtual Teams to attend a weekly tag-up about the project; move to Partners & Suppliers to check progress on parts coming in for systems test later in the week; and finish the day as member of the Public, searching for educational-oriented material on a different NASA project that he can hand out to his child's class the following day.

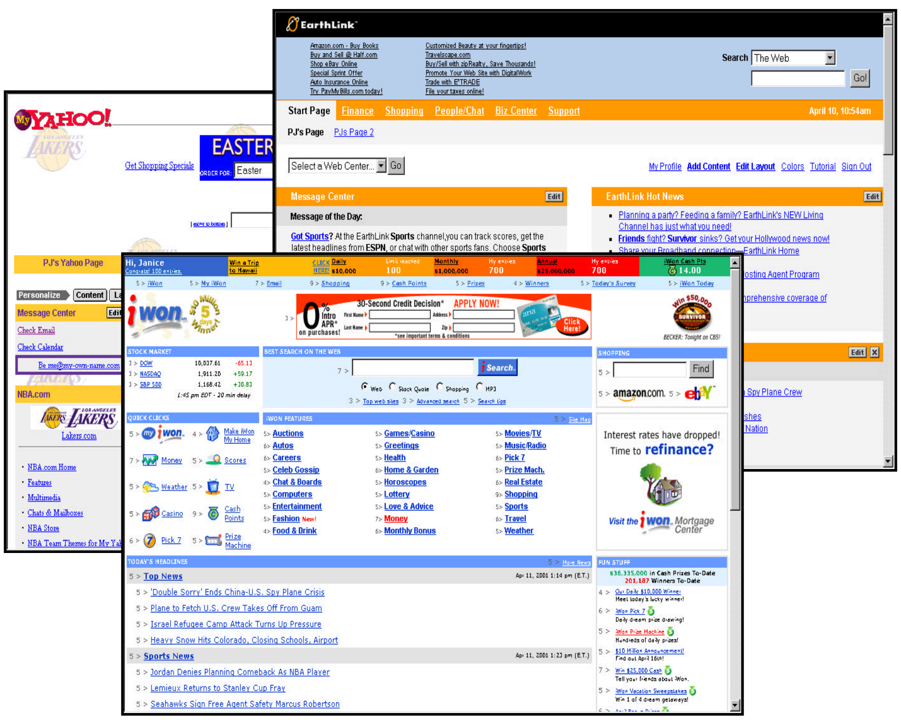

Figure 1. Examples of commercial portals include MyYahoo, MyEarthlink, and MyMSN.

A. WHY ARE PORTALS IMPORTANT TO NASA?

When fully implemented, including all the underlying technology and process efforts described in the previous section, a portal can do the adjusting among different kinds of information and different audience members:

- **Faster dissemination**. Through content management, data can be made available to all its audiences as quickly as the content-owner chooses to do so. Approvals and appropriate controls (for privacy or export control, for example) can be obtained through automated processes, saving time and, as economists phrase it, "shoe leather costs."
- **Faster search and retrieval**. Through metadata and information architecture, data can be described in a commonly accepted format. This will facilitate quick searches among the wide NASA resources for relevant information. It can even allow personalization agents to suggest to users that they look at resources they might find very useful, perhaps heading off a problem before it arises.
- **Faster information creation**. Combing information architecture with a modular system design allows new information to be easily incorporated into the portal with much less reformatting than would otherwise be required. It can also allow the portal to access and distribute information residing in legacy systems, saving the cost of perpetual reformatting.
- **Secure access**. Through access control, a user can be identified and access immediately granted to all the information to which she is entitled: financial data on her project; limited public data on other projects. This can eliminate the need for redundant control mechanisms and multiple passwords.
- **Enables a collaborative environment for knowledge management and sharing**. The ultimate objective of knowledge management is to give the power of collective expertise to each and every team member, allowing us to deliver the greatest value to our customers, the public, industry, academia, and government. In a world where speed of change, learning, and adaptation are critical touchstones of success, knowledge management is an essential capability (Figure 2). Knowledge management has to with the capture, organization, storage, and distribution of intellectual assets. It has to do with people—their learning, collaboration, and creativity. Knowledge management facilitates the reuse of proven resources and methods, reduces costly mistakes, and enables rapid absorption and diffusion of new ideas—allowing NASA and our customers to communicate more effectively.
- **Promotes customization of only relevant data to each audience**. This approach will also greatly enhance the customer experience by remembering the user's preferences and

presenting content according to those preferences on the users next visit. A user will also be allowed to set up an account with MSFC to deliver requested information, similar to a myExcite web page.

- **Decreases the possibility of out-dated information**. Calendaring allows content providers to determine when content is presented on the site and when it is removed. Automating this process ensures that date-sensitive information is available only while relevant.

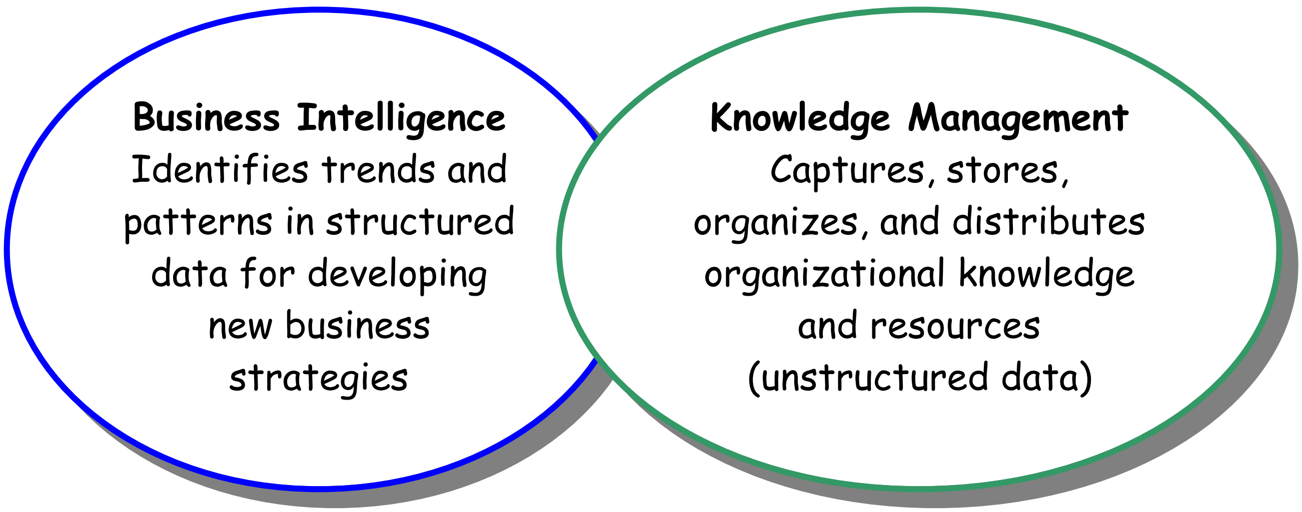

Figure 2. Enterprise information portals represent the integration of business intelligence and knowledge management technologies (adapted from Firestone, 1999).

B. WHAT DOES A PORTAL DO?

A portal is an information tool, most often using Internet technology (TCP/IP communications, web browser interface, and back-end databases) that presents a user with a single point of entry into a heterogeneous information space. Often, a portal allows users to pick and choose from a variety of information sources and “remembers” those choices upon the user's subsequent visits. Underlying an effective portal are a) creation and maintenance of relevant, interesting content; b) an efficient search tool, ideally capable of extracting information from a variety of formats and legacy systems; and c) a foundation information architecture—a system of describing the information that makes relevant data easy to find.

Some of the portal functionality can be categorized as:

- **Information Catalog Management.** Portals must provide a mechanism to organize content into categories meaningful to users (e.g., a category tree similar to Yahoo).
- **Content Management.** As portals grow in scope, content sources proliferate. Content management becomes critical, particularly since content dynamically flows into the portal in near-real time. The ability to manage this content becomes an important shared portal service.
- **Repository Management.** Portal frameworks must incorporate a repository to store information, as well as support access to information stored in file systems and other repositories (e.g., report servers, doc stores).
- **Metadata Management.** Beyond the content itself, metadata has become a critical linchpin to assist organizations in applying a taxonomy over large collections of information.
- **Personalization Engine.** Portals deliver a unique user interface by establishing customized navigational structures, content, and application interfaces.
- **User Profile/Membership Management.** Going hand-in-hand with personalization, profile management will enable users to set wallpaper backgrounds, localization parameters, and other characteristics. These personalization characteristics will be available to other applications and services plugging into the portal.
- **Activity Tracking.** Monitoring on-site behavior enables recommendation engines to suggest other information and application interactions based on explicit and implicit associations made while users traverse a site.
- **Access Control.** Authentication and access controls are essential portal underpinnings. Users should expect a single sign-on capability as they interact with the spectrum of applications, content, and services available through the portal.

C. RETURN ON INVESTMENT

The ROI to NASA will be largely realized in increased efficiencies and enhanced safety and mission success. Although dollar amounts cannot be placed on safety and mission success, the returns obviously are of paramount importance to the Agency. Although difficult to quantify, monetary savings would result from increased efficiencies and decreased duplication of effort. However, the key benefit will result from increased communication and access to data thereby enhancing safety and mission success.

If a portal allows a civil servant making $40,000 annually to spend one hour less each week searching for information and turn that time to more productive use, then over a year $1,000 in NASA payroll will have been put to better use. While this does not constitute direct cost savings, if a portal can cause similar shifts to occur across the Agency's civil service and contractor work force, NASA will be redirecting hundreds of thousands or even millions of dollars annually from unproductive activity to substantive work.

The expected ROI for a portal could occur in as little as 1 year of operation.

D. CURRENT NASA PORTAL EFFORTS

There are several portal efforts underway or in operation at the Agency. This paper looks at the integration, specifically, of those called MyNASA, InsideNASA, and the Mission PI. More globally, efforts have been made to look at the SAP technologies available through IFMP

The intent is to provide a framework for other portal developers to build upon, while pushing for interoperability with existing or developing portal activities (such as the Technology portal from Code R and the efforts in IFMP) (Figure 3).

Figure 3. Concepts of Agency-wide portals, and NASA portals in actual use, give a flavor of what InsideNASA and MyNASA could look like.
Technology Portal: **http://nasatechnology.nasa.gov/index.cfm**
InsideNASA: **http://eis.jpl.nasa.gov/~jedutra/insidenasa** *and* **http://www1.msfc.nasa.gov/ENASA/**
PBMA Knowledge Management System **http://pbma.hq.nasa.gov/pbmamaster.html**

2. REQUIREMENTS

The requirements for a NASA portal come from many sources. This report references the requirements documentation for portal selection of for InsideNASA, MyNASA, IFMP, MSFC content management, and Mission PI activities and has integrated those documents into a single matrix (see the Appendix).

At the highest level, the requirements for a NASA-wide portal for internal and external users can be looked at as fulfilling the need to:

- Provide tools and strategies to manage our online information environment and help NASA employees and partners meet new mission challenges and retain our organizational knowledge
- Support online communities to encourage knowledge sharing across disciplines
- Easily and efficiently integrate existing web-based resources to support easier access to information through search and browsing
- Deliver personalized information (both through subscription and enhanced access)
- Offer fast, accurate, and simple search for general public surfing Agency public web sites
- Provide a deep, broad, complex, intelligent, and personalized search options for partners and power users to search across the Agency's repositories
- Deliver more rapidly to the NASA environment commercially available tools that will put the full power of the Web in the hands of NASA personnel

The Appendix shows a requirements matrix for both the InsideNASA and MyNASA portals, as well as initial prioritization of those requirements for the phased delivery. Requirements will continue to be gathered throughout the initial start up of the project. At the time of the Requirements Review, the requirements for each phase will be baselined.

3. SYSTEM ARCHITECTURE

A. DESIGN PRINCIPLES

Certain principles should guide the design of the portal. As the underlying technologies might change, these principles should be upheld to ensure that future deliverables meet the goals and best practices initially set for the portal. These principles (Figure 4) are

- ***User-Centered Design***: Information must be presented at appropriate level for target audience and usability will be a key design and testing criteria.
- ***Platform Independence***: End user devices shall be platform, operating system, and browser independent. Current end user devices include Windows, Mac, Unix, Palm, and cell phones.
- ***Open***: Users will be distributed across the globe and include NASA employees, external partners, other Government agencies, international partners, and the public.
- ***Modular***: Components from multiple vendors can be swapped in/out; external content and applications will plug in.
- ***Leadership***: Promote a common portal approach across the Agency.
- ***Security***: All users will be authenticated and assigned into one or more groups based on their authorized role(s); information must be protected against unauthorized access or modification.
- ***Standards Based***: Leverage ubiquitous Internet standards and directions.
- ***Integrated with NASA Infrastructure***: Build to easily integrate into existing Center architectures and policies.

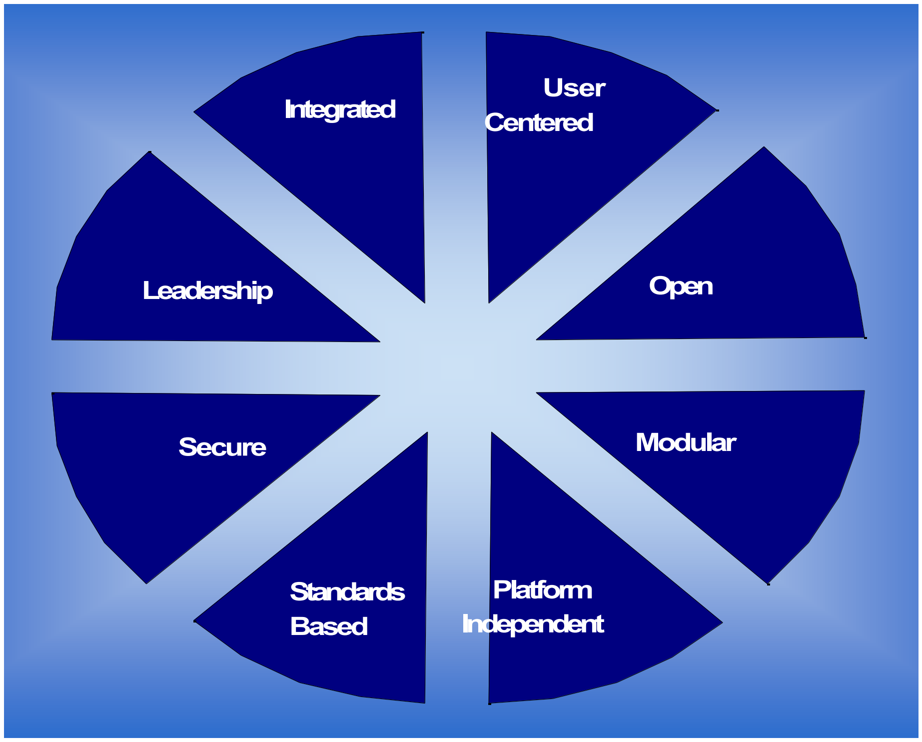

Figure 4. Portal design principles

B. USER INTERFACE

In applying these principles to the portal's user interface, it is critical to reinforce this portal as the user's single entry point. Each user can customize the content and layout to maximize their own efficiency; however, the goals for the generic user interface should include:

- Follow *NASA Web Best Practices*[4] as appropriate
- Provide a consistent look and feel, including the use of the NASA logo, privacy statement, search function, and contact information
- Interface design is about visual guidance. How navigation options are presented is closely tied to how usable they are. If they are hidden, difficult to find, look too much like text, or are otherwise visually confusing, users will have trouble navigating.
- Support easy and efficient navigation
- Be consistent in the placement and design of navigation elements for the generic portal. Users have the right to expect navigation buttons and bars to show up in the

[4] *NASA Web Best Practices*, December 2000, http://nasa-wbp.larc.nasa.gov/

same place every time. Consistency builds the user's trust and enhances the quality of the experience.

- Do not violate expected browser behaviors that users have come to know and trust. For instance, they have come to expect that their browser's back button will always work a certain way and to break this rule without a significant reason is a breach of usability.
- Provide a Home button on every internal page
- Provide overview and frequently asked question (FAQ) pages to give the user background information on the technology and their choices within the system
- Provide e-mail contacts and other points of reference for additional help (such as a Help Desk)
- Provide online help that is contextual and easily available
- Usability testing should be conducted to ensure ease of task completion and resolve any outstanding issues related to key points in portal functionality
- Users should be able to complete required tasks easily. These might include editing data channel content and layout

C. HARDWARE, SOFTWARE, AND NETWORK ARCHITECTURE

Portals can be divided into four layers as depicted in Figure 5. NASA should take an architectural view of portals by defining services at each layer, rather then focusing on the purchase of a single vendor monolithic product to meet all of our requirements.

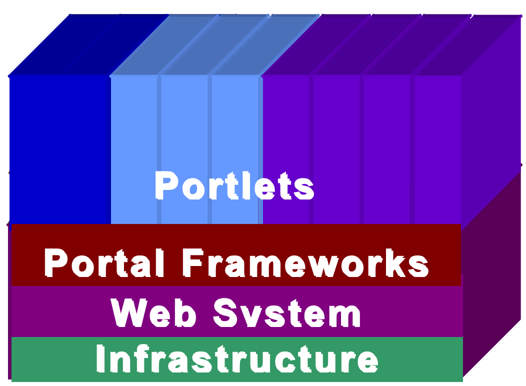

Figure 5. Portal architecture (taken from MetaGroup)

1. Layer 1: Infrastructure

Common infrastructure layer components required to enable a portal are: networking, directory services, email, security (authorization and authentication), end user devices, databases, and enterprise application integration (EAI) services. The NASA networking service, email, and end user devices are ready to support a NASA portal today. Other areas of the NASA infrastructure are not necessarily ready to support a portal, in particular NASA does not have an application-ready corporate directory, nor do we have a fully deployed authentication service that can serve all NASA employees, external partners, and public users.

The portal should be attached to a dedicated high-speed network (e.g. switched Gigabit Ethernet) to ensure optimum performance. Connectivity to the NASA network will be dependent on whether the portal is hosted in an external outsourced environment or at a NASA Center. Assuming external hosting, connectivity to NASA networks will be via in-place Internet connections.

End user device support will include all devices and software defined in NASA STD 2804/2805. Further, flexibility to support emerging mobile devices such as PalmOS and cell phone systems is highly desirable.

Authentication will be enabled by a best practice capability already deployed to thousands of NASA employees within several centers and mission areas. The capability provides two-factor authentication using a hardware token and is based on the RSA SecurID product. The authentication service will be provided by the Secure Nomadic Access (SNA) project.

NASA does not have an application-ready corporate directory service, so for this functionality the portal will use the directory and or membership services provided within the portal framework.

2. Layer 2: Web System Services

The web system services layer is a specific part of the infrastructure that pertains solely to the Web, and is a key enabler of portal systems. Web system services consist of business logic management and data repository access—or what is commonly referred to as “application server” software. Web servers are another key component of the web system services layer. NASA does not have any preferred corporate application server or web server software at this time. This is expected to be addressed by the Web Management Team.

3. Layer 3: Portal Framework

The portal framework is commonly packaged into a commercial "portal" product and consists of components for personalization, profiling, profile management, metadata/taxonomy, content management, access control, and activity tracking. In some cases, some of these components are done by external best of breed services, the most common example being content management.

The portal framework includes an API for plugging external components into it. Many portals also include connectors for easy integration of popular applications in categories such as enterprise resource planning (ERP) and collaboration. Portal application integration capabilities are beginning to overlap with enterprise application integration (EAI) functions (EAI is a separate class of vendor products).

4. Layer 4: Portlets

Portlets are external applications and repositories that you want to link into the portal. External applications that may be desirable to link into a NASA portal include: email systems, collaboration systems, project management systems, content repositories, and ERP systems (e.g. IFMP modules). EAI style connectors are beginning to be seen in portal framework products, enabling integration with portlets.

Based on this discussion, we can see a high-level architecture for the NASA portal as depicted in Figure 6.

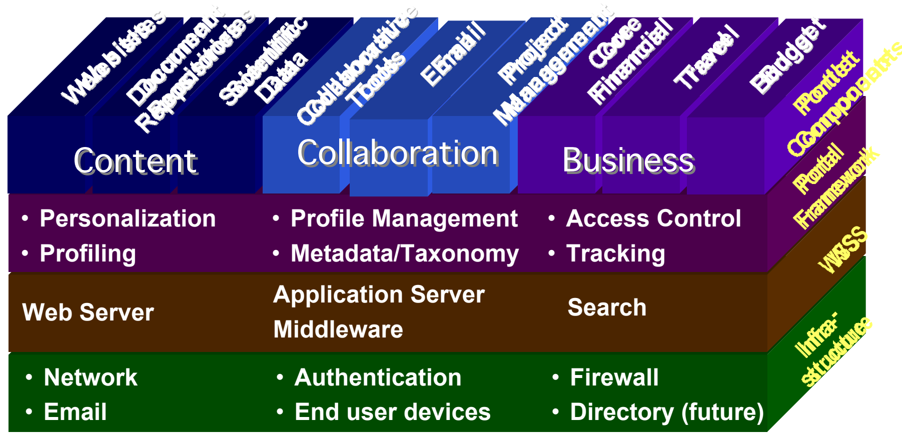

Figure 6. NASA portal architecture

D. SECURITY SYSTEM AND ACCESS CONTROLS

There are many documented threats to our information systems that must be mitigated by policies, processes, and technology. The portal will follow NPG 2810, which establishes basic IT security requirements for NASA systems. In particular, the NASA portal must protect against:

- Unauthorized access to information or applications
- Unauthorized modification (or defacement) of information
- Denial or degradation of service to customers

To mitigate these threats, the NASA portal must implement the following controls:

- Authentication for all users
    - NASA employees via a two-factor (or strong) method
    - A preference for two-factor (or strong) method for all other users when possible
    - NASA external partners via NPG 2810 compliant passwords
    - Public visitors via simple passwords (chosen by user)
- Role-based access to information
    - Users will be authorized into one or more roles
    - Each role will form a group of users
    - Example roles could include
        - NASA employee
        - NASA management
        - Support service contractor
        - External partner
        - Member of project X
        - Public (student, media, etc.)
- Information access control
    - All information objects will have access control lists (of groups and/or individuals)
    - Read/write/delete object privileges will be available
- Information security considerations will be part of portal information hierarchy and taxonomy design so that access control is enabled
- Firewall

E. STANDARDS

There is no generic portal API or other standards in place in the portal industry and none are expected in the next 2 years. However, there are standards we recommend for:

- Connecting portal to infrastructure components include: LDAP (directory access), POP, and IMAP (email store access), and HTTP (end user device access).
- Connecting portal framework to portlets include WebDAV (document/content management access) and enterprise application integration (EAI) products.
- XML can be used in several areas such as the portal API and metadata management, but payloads are proprietary. XML can also be used for access to the emerging area of "Web Services," which is a method for applications to expose coarse-grained services via industry standard protocols (e.g., SOAP and UDDI). Web services provide system-to-system communications only (not for presentation).

We have a preference for a Java programming environment (de facto standard) due to its platform portability.

4. INFORMATION ARCHITECTURE

Internet users have historically employed two approaches to finding information on the World Wide Web: 1) browsing across topical subject categories, and 2) searching for specific instances of informational keywords. Both these behaviors need to be addressed in order to give users support across the full range of information navigation services.

A. INFORMATION ARCHITECTURE: WHAT'S BEHIND THE DOOR?

A top goal of any portal is to act as a single point of access for users to find desired information generated by the enterprise on topics of particular interest to them. The portal is literally a door to an Agency information catalog rich in knowledge about the topics in which NASA specializes. NASA discoveries and research over the decades span a wealth of information in technology, planetary science, engineering, and many more subject areas. Information architecture seeks to create a topical framework that embodies and enables these areas of interest.

For the purposes of this paper, we will use the definition of information architecture first coined by Richard Saul Wurman in the 1980's

Information architecture is the art and science of structuring and organizing information systems to help people achieve their goals.

For this report, we will focus on the aspect of information architecture (IA) that concerns itself with designing organizational systems for content, creating consistent labeling schemes, and devising navigational pathways through sets of associated data. (This is only a portion of the larger IA that NASA needs to address.) The goal of information architecture is to facilitate knowledge access by building taxonomies, categorizing information, and creating site maps to enable user exposure to relevant material. In order to accomplish this, we must first understand our content and users.

As we examine the contents of NASA web space, it is useful to know the audience types that will be visiting the site and what their needs are. The NASA portal will be developed for two primary audiences: internal and external. Clearly, the needs of these two audiences and their many communities are different and, therefore, portal navigational mechanisms should reflect their distinctive requirements.

External user groups break down into communities. A report was recently prepared for Code P by SAIC, analyzing the external audience demographics and likely visitors to the **www.nasa.gov** site and any future public-facing portal.[5]

- Internet User Type A
    - Families with children (age six and up)
    - Students
    - Large institutions
    - Some large corporations
- Internet User Type B
    - Business professionals
    - Academics
    - Engineers
    - High school and college students
    - Government administrations
- Internet User Type C
    - Space enthusiasts
    - Scientists
    - Media
    - Researchers
    - Computer enthusiasts

Internal audiences may look very different. They may be organized by job family or by their role on a mission. Research needs to be done in order to properly define and characterize the base customer groupings. Many organizations are engaged in studies that profile their core user groups. They are detailing typical tasks that users regularly perform using the Web and identifying information repositories that must be accessed in order to complete those tasks. This type of information leads to a better understanding of the needs and requirements of the users.

Once there is a common understanding of our customer base, we can begin to address content classification and information architecture issues. A strategy to formulate a useful

[5] Pino, Chris and Brian Dunbar, *HTML 4.01 Tools, Preliminary NASA.gov web Audience Profile*, SAIC for Code P, NASA HQ, National Aeronautics and Space Administration, Washington DC, May 2001.

underlying architecture for NASA web space can be developed from the consideration of three main components (Figure 7):

- Content
- Business context
- User feedback loop for continuous improvement

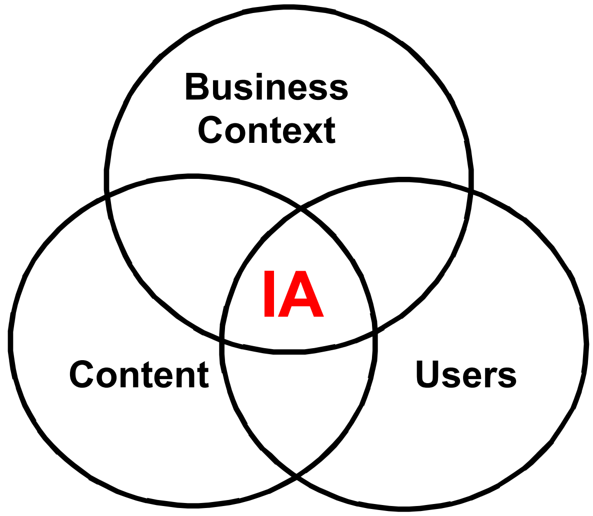

Figure 7. An Ecological Approach to Information Architecture

B. GENERAL STRATEGY: TOP DOWN AND BOTTOM UP

NASA web space has many pre-existing sites on a plethora of subjects. With a *top-down approach* to topical organization, a main hierarchy can be determined that offers navigational pathways for users to take through a NASA Web space directory. This directional approach emphasizes a broad view and includes large topic groupings through which the user can drill down to desired content areas.

With *bottom-up information architecture*, individual content chunks are considered and bridged to the site through navigation from the lower to the higher levels (Figure 8). There are different levels of granularity of information architecture for any site. An information space as large as NASA's will take some time to analyze and categorize.

Both methods work together to create a cohesive web environment. Most sites are a combination of the two, though some sites are more focused on one than the other. Good information architecture is invisible if done well.

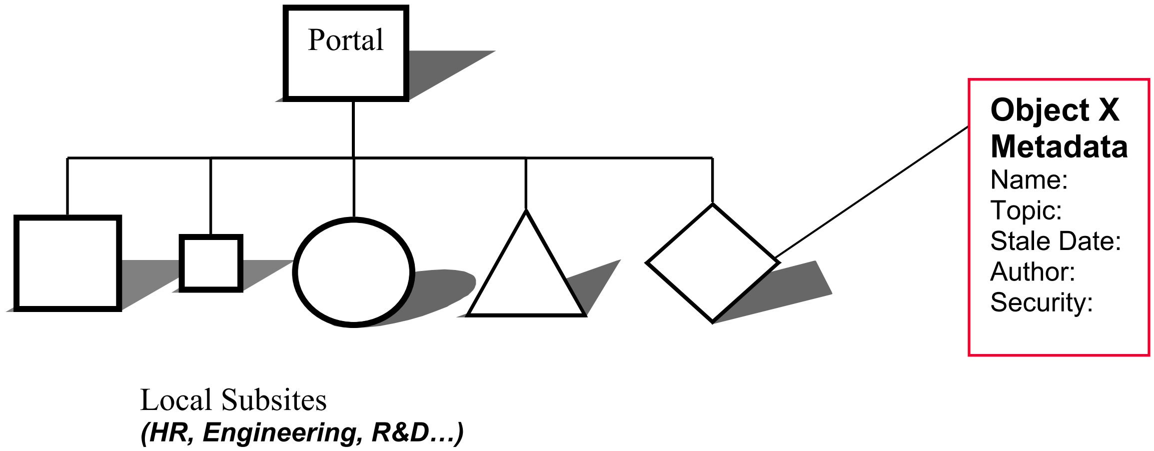

Figure 8. Information architecture from top to bottom

C. ORGANIZING CONTENT AND CLASSIFICATION SCHEMA (TAXONOMIES)

After analyzing user needs, an initial step in designing information architecture is to carefully consider the content of a site. A content inventory and analysis should yield a clear understanding of content requirements. A content map is developed to facilitate the organization of content into intuitive groupings for user browsing.

Once the content is well understood, study can begin on a classification schema that fully describes electronic assets. The general trend in data architecture systems design is in a deconstructionist direction, breaking content down into information objects. Taxonomies expose relationships between data objects and provide a blueprint for an integrated view into the information space. In other words, once the building bocks are broken apart, they are ready to be glued back together in ways that reflect user understanding of the information environment.

A key part of information architecture is the design of taxonomies that introduce users to related ideas. Users may all see the same information; however, they traverse through it in individual ways that reflect their experience and need to know. The associations one makes provide a creative experience and enlarge the value of the NASA web and its possible knowledge discovery connections. Taxonomies are information access tools that encourage brainstorming, collaboration, and improved communication.

Furthermore, taxonomies are necessary due to the complexity and subtlety inherent in language and information retrieval. Most people insert keywords into a search box and click the "Go" button. Keyword searching assumes that individual terms line up with concepts. But, language is inherently ambiguous, so keyword searching often ends in user frustration. Thus, *taxonomies create a contextual framework for information retrieval while mitigating the complexities of language*.

But how do we go about creating taxonomies? Once the content map is created as described above, the next steps are an examination of the content looking for patterns and relationships within the material. The architect is seeking content attributes relevant to users. Content groupings tend to gravitate towards a natural state of affinity. Some classification schemes are known as exact schemes, such as categories that are grouped in the ways below.

- Chronological
- Geographical
- Alphabetical

Other schemes are known as ambiguous schemes. Some examples are:

- Subject
- Audience
- Task

Ambiguous schemes usually provide for more than one entry into the resulting directories, which gives the user a better chance of finding the desired term. Providing multiple pathways to information is usually a good idea since it increases the chance that users will find relevant content and improves usability.

In general, taxonomies progress from the "genus to species" classification model, meaning that they go from broad categories to more narrowly defined groupings. It is important to remember that there is not always one way to express a "best" taxonomy. Taxonomies can be cross-faceted to express attributes of significance to a varied audience. This allows for a flexibility and robustness to site directories and allows them to address the needs of a mixed audience that may have differing goals when visiting a site.

D. BUSINESS CONTEXT AND CLASSIFICATION SCHEMA

Once a first draft of taxonomies is formulated, it is time to ask content providers, site designers, and other stakeholders if the taxonomy adequately describes their goals for searching and finding online information. In this way, we can check for the validity of the taxonomy in relation to the Agency's desired goals for the portal.

Other goals may include better information sharing among groups such as communities of practice or project teams. The IA developers should study the design of targeted taxonomies tailored to specific team needs. The NASA thesaurus may be of use to identify commonly referenced terms and popular topics.

It is important at this point to consider the scalability of the architecture. Will it suffice in one year, three years, or further? Since the underlying information architecture will be used as a framework to build or map associated metadata, we want to be sure the taxonomy is stabilized and well tested before we begin the next step in developing content management systems. However, taxonomies represent our understanding of the world around us and, as our knowledge evolves, our navigational signposts will have to change as well. So we need to consider long-term maintenance and care by information stewards for the classification schema.

In addition, we will want to consider any upcoming initiatives, branding strategies and NASA's placement in a larger Federal information architecture, such as FirstGov or other e-government projects.

E. WEB DIRECTORIES AND USERS

Once in place, testing of the taxonomy should occur to see if it is adequately serving site users. Refined taxonomies are generally known as directories. The descriptive words used in directories are extremely important to the ease of information access. Usability testing can confirm which areas are performing at designed levels and which areas need to be re-evaluated.

Jared Spool and Erik Ojakaar discuss typical information foraging behavior on the part of users. They advocate the development of practical taxonomies and "trigger words" that users will typically employ to find certain site material. They recommend the study of user click streams to better understand how users think about site content. They also suggest that the most effective directory structures are designed to expose subcategories, thereby giving the user more clues as to what they can expect to find at the other end of the link.

In addition, they recommend a careful examination of site search logs. Many times, user queries gleaned from these logs can expose what users are really looking for when they come to a site and what words they use to describe the information they seek. Search log analysis can often provide information architects with valuable clues to user needs and typical behaviors.

Spool and Ojakaar also cite the notion of a "level of agreement" from users about what certain terms mean. The more universally accepted the term, the more successful users will be at following the logic of the directory entry to where content resides. Therefore, the more descriptive the terms used in a categorical grouping of site content, the more "intuitive" the site is said to be. This underscores the importance of utilizing users' preferred cultural language and vocabulary terms when it comes to directory design.

Taxonomies are most often developed through an iterative process, meaning that they may take several generations to evolve into their most effective version. In addition to describing the existing corpus of material, new content becomes available and users' needs may grow in unexpected ways. Therefore, it is the recommendation of this team that work on the taxonomies precedes the design and implementation of complex content management systems.

F. THE ROLE OF METADATA

Once a stable taxonomy has been developed that meets the needs of users, we can begin to build metadata tables that express content attributes. Metadata definitions will generally develop into XML-derived solutions for content management and reuse further down the road, so they are a key piece of the information architecture to get right. SOAP, RSS, RDF, ebXML; these are all specifications that are increasingly coming into common use. In order to fully leverage the power of data interchange, we will want to be sure we have a solid foundation to build on. Metadata represents the building blocks of that foundation.

Metadata can also be optimized in a number of different ways. Metadata attributes can address the varying needs of audience and allow for flexibility within a schema. Metadata attributes that are related but include characteristics of importance to different audiences are said to be cross-faceted—they can serve several audiences at once.

Metadata can be used for helping to implement publication-related business rules. For example, many content management systems use an "expiration date" tag to remove content from the Web that is no longer considered relevant. In the same way, metadata could be used to mark a document with an "Archived" stamp, indicating to users that the information they are viewing is no longer current and should be treated as such when making engineering, design, or business decisions.

Metadata can be used to help enforce authorization business rules. Documents tagged with different levels of classified status can automatically allow a certain class of users to view them and turn away others. This would allow certain external partners (such as academic researchers) access to all public information and only that internal information to which they have been granted specific access.

It is possible that different communities will want to select tags they feel best describes their content. In that case, it would be appropriate to consider the creation of a metadata registry at the Agency level that includes an underlying thesaurus for the reconciliation of differing metadata structures. This would allow communities to control their information space at the local level, while integrating their specific constructs at the Agency level. This is an example of the top-down and bottom-up strategy described earlier.

Although we want to allow for the heterogeneous expression of data in ways that communities find intuitive, we also want to encourage the adoption of standard tags developed from an existing set, such as the Dublin Core metadata specification. A small set of commonly used core tags should be recommended to all Agency Webmasters for use on web materials.

In order to get the most out of the effort to tag web content consistently, first-level controlled vocabularies could be developed that include some simple synonym lists most commonly used by NASA personnel. Hence, mission names or Center names that are commonly abbreviated should be included in keyword descriptions.

In addition to supporting the foundation for content management systems, metadata facilitates keyword searching. Because site users utilize search extensively, we want to be sure that standards are known and implemented across the Agency by NASA Webmasters. This will take some time and resources dedicated to education and training. The NASA Webmasters Group can be instrumental in providing support in this area. Standards can also be reinforced by using the web management model being developed by the Web Management Services Team.

G. INFORMATION ARCHITECTURE AND DATA ARCHITECTURE

A discussion about metadata would not be complete without a side bar on data architecture and how the two relate to each other. The underlying data infrastructure system is critical to how information is passed over the Web. Data must have mechanisms to move through cyberspace and to specified destinations. Part of the infrastructure is expressed by coded middleware wrappers using XML to tag the data with identification characteristics and affiliated uses.

In addition, new tools such as Web Services Description Language (WSDL) and Universal Description, Discovery, and Integration (UDDI) are becoming available to extend the power of data architecture. It will be necessary at some point in the future to undertake a study of NASA's existing data architecture model and how it can be extended to handle increased traffic flow.

Data architecture also addresses problems of interoperability between systems that have been developed separately. It provides ways for information to move across the Web between providers and consumers of needed data. The technology components of a mature data architecture compliment the information architecture and allow it to function at its peak.

H. INFORMATION ARCHITECTURE AND THE USER INTERFACE

The art of labeling is an often overlooked component of information architecture and yet, since it defines the interface between the user and the content, it is one of the most important aspects of site organization. Labeling systems are expressed in site navigation mechanisms such as tabs and buttons, so care should be taken to see that they are thoughtfully developed, easily understood and consistently applied.

Once top-level Agency constructs are in place to give the web environment some definition, content publishers from different parts of the Agency can begin to see themselves as a part of a larger community. They can then better understand the benefit of employing language that is universally understood in their content and that its use maximizes their interface to NASA users through NASA portal data channels.

The user interface often impacts how visitors interact with a site. Information architects create wireframe models of web pages that map out functionality and navigational pathways. These wireframes are employed in testing to see if the material is presented in a framework that enables site comprehension by the user and promotes usability. This not to be confused with graphic presentation, but rather expresses a visualization of how content is organized and presented to the user. It is meant to diagram the site's underlying information flow and express the various ways that users traverse the information space.

I. INFORMATION ARCHITECTURE AND KNOWLEDGE MANAGEMENT

Knowledge management echoes the concerns of information architecture in its desire to aggregate and order the intellectual assets of an organization. Both disciplines seek to promote greater efficiency and productivity through better management of content crucial to enterprise success. Knowledge Management seeks to encourage community interactions and knowledge discovery through the creation of common platforms that house enterprise information sharing systems. As users browse through structures of knowledge, they refine and extend what they want to know.

By defining a NASA institutional information architecture, knowledge management's goal is to determine the scope and landscape of the NASA web domain. In making the web environment an easily accessible resource, the team is enabling innovation and knowledge reuse as well as speeding the process of placing the user as close as possible to desired information.

J. SUMMARY RECOMMENDATIONS

The steps to develop an effective information architecture for NASA web space can be summarized as the following:

1. Identification of audience types and needs – internal vs. external
2. Identification of top NASA goals for NASA web space
3. Inventory existing content
4. Content requirements analysis
5. Mapping of content to user needs and creation of information architecture blue prints
6. Development of topical taxonomies from blue prints
7. Validation of taxonomy alignment with Agency goals
8. Metadata development and iterative user testing
9. Descriptive labeling systems consistently applied across NASA web space
10. Processes established for ongoing identification of documents with metadata attributes
11. Testing for continuous improvement

The long-term methodology for developing a robust information architecture for NASA is mapped out above. It would involve a team to do research on audience types and perform a content inventory. This might be facilitated by support from the eNASA Web Services Team as well as the NASA Webmasters and take about two to three months. Content analysis and mapping could take longer depending on the scope of the effort and the resources available.

Taxonomy design evolves from content maps and usually requires expertise in classification methodology. It would be helpful to engage a professional information architect with library training from an association such as the American Society for Information Science and Technology (ASIS) to suggest groupings that reflect the best usability. The time frame for development of metadata tables depends on the scope of the project, but typically range from 3 to 6 months. The acceptability of associated metadata is usually more of political issue than a strictly IA issue. Getting buy in and agreement from all parties may take time. Implementation of tags will take more time as Webmasters will probably have to procure funding in order to do the necessary HTML production work. Financial support and support from upper management will determine how quickly this can be accomplished.

K. GOING FORWARD: EVOLUTIONARY INFORMATION ARCHITECTURE

As NASA moves forward in its transition to a web-enabled organization, information architectures that provide scaleable frameworks for web assets are an indispensable aid to navigation by both the internal and external users. The Web Management Services Team is studying ways to better manage web publication processes. In the future, we will have to consider online web services and applications as well as content integration into our directories.

Internal and external directories will overlap, but the needs of the audience bases remain quite different. Internal groups may be centered on communities of practice that reflect technical disciplines, such as thermal engineering. Other groups may want to use the portal as a platform for collaborative partnering. Most NASA missions are spread across Centers and the world, and the portal can facilitate work for teams that are not co-located. In the future, it is possible that these groups may build their own directories that point to web assets that are of particular value to them.

The Web is a powerful communications medium, capable of carrying many types of information. In order to better structure NASA's information space and facilitate retrieval of relevant data, it is helpful to provide a foundation of information architecture. Information architecture acts as a framework for users and helps them develop a mental model of how online assets are organized (Figure 9). This, in turn, stimulates more successful interactions between users and the NASA web, improving the quality of work performed by NASA employees and effectively communicating the goals of NASA missions to the public.

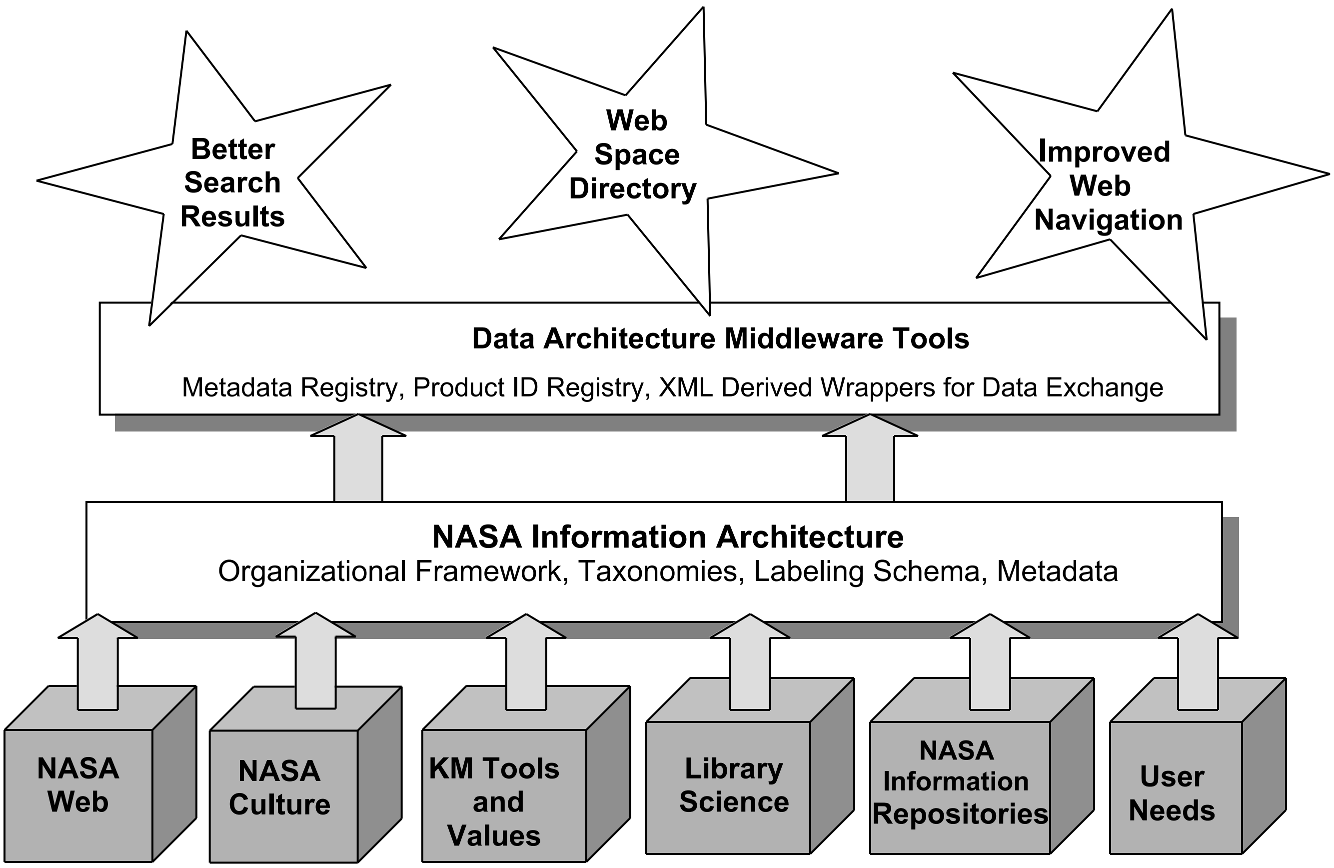

Figure 9. Building blocks and information flow of NASA information architecture

5. CONTENT MANAGEMENT AND BUSINESS PROCESSES

Content management describes how the creation of information is facilitated, and then stored and delivered to the user. The theory behind a successful content management architecture is not just in having a good technology but also in having the right business processes that takes the fullest advantage of a content management system. Basically, content production for a portal (and really any website) breaks down into two components: the processes by which content is created and the technological systems that facilitate its distribution. Both are critical and are as equally important. Without the right processes the technology will not be much more than some expensive pieces of hardware and software, and without the technology to facilitate, automate, and standardize the business processes, there will be no good methodology to bring together the myriad of processes that exist within NASA.

In the current NASA environment, content management often rests entirely with the content owner. He or she creates, manages, distributes, and maintains the content in near complete autonomy of all aspects, from factual veracity to graphic presentation. In a portal environment, content management becomes a more cooperative process. The content owner still creates the content and starts the publication process, but other aspects of publication, such as policy compliance or graphic presentation, may rest with a different person or group

A. BUSINESS PROCESSES

Part of the creation of a successful portal has little to do with the actual technology. Defining the business process by which content will be created, vetted, and approved for intended audiences is critical to managing the content. The business processes used to publish information at NASA are as diverse as the elements within the Agency. For example, NASA's Scientific and Technical Information group has long-standing processes that have served the Agency well, but which are different from the Office of Public Affairs' equally long-standing and effective processes.

Establishing business practices to be incorporated into a web/portal services model is the purview of the Web Services SRR team. Here it is appropriate to recommend a few operating principles. Before doing so, it might help to clarify how two specific terms are being used.

- Content creation -- the creation, compilation and rough organization of text and multimedia related to the topic at hand.
- Content management -- the formatting, whether for the Internet or other distribution media, workflow, distribution, and archiving of the material.

Principles

1) Content creation should remain with the owners of the content. They are the best qualified to know what kinds of information their audiences need and to ensure its accuracy and timeliness.

2) As a corollary, web/portal business processes should not include content review, but should rely on existing review processes. Material that has already been cleared through the creating office, such as press releases, should not have to be cleared by a web/portal review board.

3) Web/portal review should be of process, so that the Web Services Team, however that may be defined:

- Knows who the content owner is
- Knows the appropriate points of contact
- Has an agreement with content owners that covers:
    - Who will perform policy (e.g., Section 508, COPPA, privacy) and security reviews—the content owner, the web services group or other group responsible for a specific function
    - The kinds of content that will be provided
    - The intended audience
    - Frequency of updates
    - Relation of content to established information architecture
    - Retirement/archiving plans

4) Depending on how web/portal services are implemented across NASA, content management may also reside with the content creators. If that is the case, the agreement described in section 3 above should be expanded to cover topics such as NASA Webmasters Best Practices.

B. CONTENT MANAGEMENT SYSTEMS

Because there are so many different business processes for publishing in the Agency, it will be necessary to implement a technology that will facilitate and integrate these processes to present a unified front to the portal audiences. Several unique processes were mentioned above, such as those in Public Affairs, STI, content creation, and web/portal teams. The right kind of technology, through hosting site components in an object-oriented, database-driven architecture, will be able to incorporate all these disparate processes into the same system so that NASA can

- Make effective use of all internal resources
- Reduce time required to implement site changes, or re-design a site
- Ensure the availability of timely, accurate information
- Scale web site to keep pace with organizational growth

A typical content management system hardware configuration may look like the one shown in Figure 10. From the hardware perspective, it is important to evaluate potential systems to ensure compatibility with existing technologies and future requirements. Some things to consider: What client and server operating systems does it support? Is it scalable? How is the usability and manageability? What about security management and replication capabilities? What web application servers and development tools are supported? Does it integrate with legacy systems?

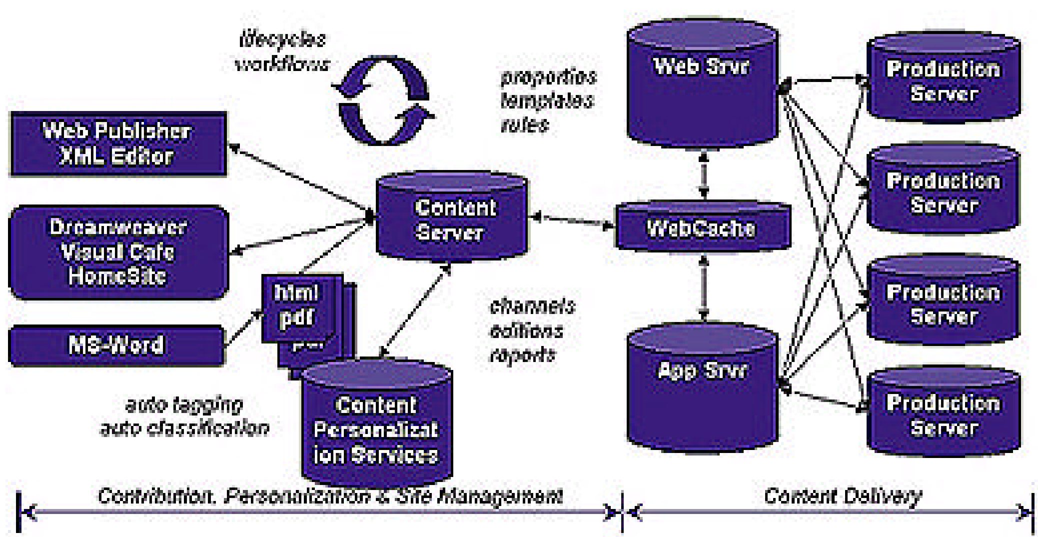

Figure 10. Content management hardware configuration (sample)

A typical content management system software configuration may look like the one shown in Figure 11. From a software perspective, it is important to evaluated functionality such as authoring capabilities, template creation, workflow, library services, access services, delivery services, personalization, and site management. For the most flexibility, the system should be modular, so that it can be scaled up or down when needed.

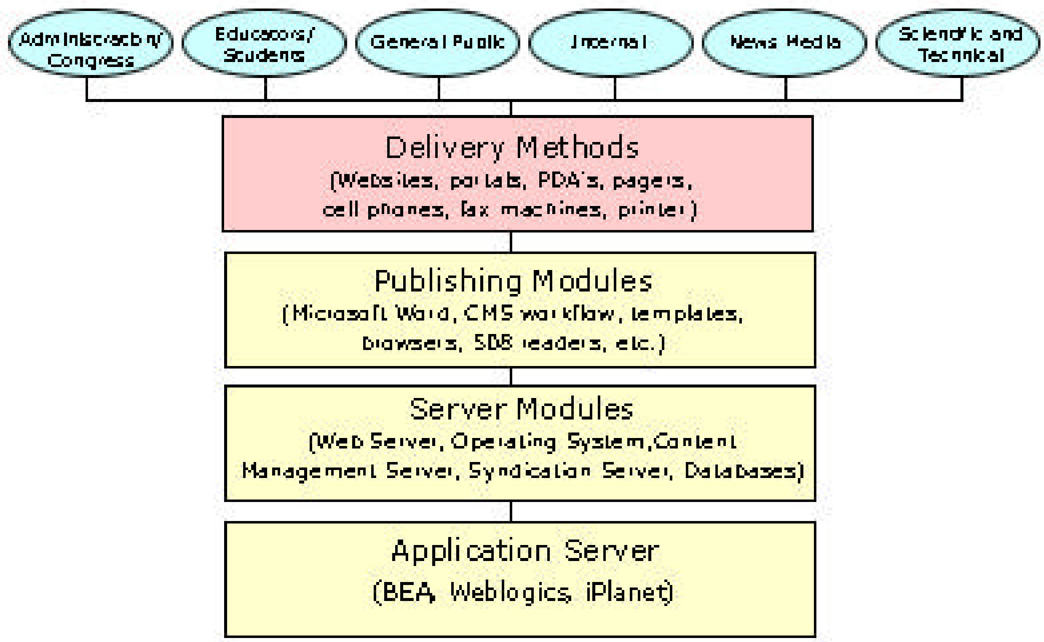

Figure 11. Content management system software configuration (sample)

A content management system encompasses a component-based architecture that separates content from presentation format and dynamic serving of content. Personalization of web content requires that content and presentation are not intertwined. That way the same information can be served up in various formats, depending upon the site users needs and interests. This capability, coupled with a page serving model that builds pages on-the-fly as they are requested, provides the foundation for delivering personalized content tailored to the user' s interests and/or browsing device (Figure 12).

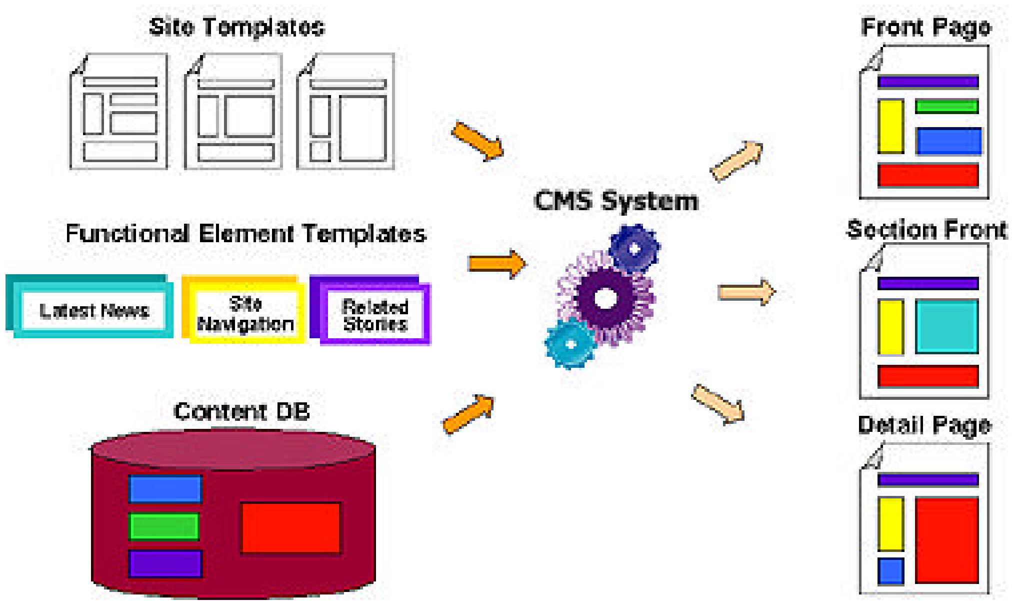

Figure 12. How content management can improve web publishing processes

Workflow - Improved business processes by automating. A workflow can be defined according to review and approval policies for each individual department or group to automatically route new web content through its approval cycle, decreasing the time it takes to get content approval.

Non-technical users can update site without affecting presentation - Decreased on-going maintenance costs. Once the initial structure of the site is built, maintaining it is done in a web form, making knowledge of html unnecessary. A content provider will simply fill out a web form with the appropriate information and hit "submit" to begin the approval workflow. The content provider will establish the posting date when inputting the information. Once the content has been routed and approved by all, it will automatically be posted on the appropriate day.

Reuse of content - improves data integrity and reduces server space needed. In many cases, more than one website will use the same information, but re-programmed to fit their own site design. A database-driven site stores individual site elements as objects in a database, pulling it together with a script into a template. With this structure, many templates can pull the same content object into their site, but it only is stored in the database once. This not only ensures data integrity, it also saves storage space.

C. TECHNICAL EVALUATION

A great deal of research has already taken place within centers and the Agency regarding vendors and products for content management technologies. However, it should be the responsibility of the Agency web Management group, once formed, to make the final decision for the Agency regarding a Content Management architecture. When deciding on a product the Web Management group should take, at least some of the following criteria into consideration:

- Functionality - Authoring Capabilities, Templates, Workflow, Library Services, Access Services, Delivery Services, Personalization, Site Management
- Technical Architecture - Client and Server Operating Systems, Scalability, Usability, Manageability, Security Management, Replication Capabilities, Application, Development Tools, Integration
- Cost - Initial Cost, Maintenance Cost
- Vendor Services - Professional Services (project management, consulting, integration, training); Support Services (availability, method of delivery, quality and responsiveness)
- Vendor Viability - Financial viability (based on financial performance over the last 3 years), Organizational Viability (how stable is their management structure), Market Viability (how strong are they in the government market).
- Vendor Vision - How vendor will evolve to incorporate new functionality into their product; Service Vision (support of their client moving forward); Company vision (What is their 3-5 year vision).

6. PORTAL MANAGEMENT

A. HOSTING

It is reasonable to expect that the first instance of InsideNASA may be implemented and managed at a NASA center. This will allow the responsible internal integrators, developers and content providers a friendly environment supported by existing infrastructure organizations with which they already have close working relationships. . Once the general confidence factor is high enough (through testing, use and rapid feedback) to recommend a solid baseline and the detailed process flow for content creation and delivery are well understood and integrated into existing NASA processes, the system could be promoted into a commercial managed hosting environment. This will allow us to focus on core competences and mission-critical processes rather than dedicating personnel, resources and activities to web hosting of NASA Portals.

The growing importance of e-business and the increased demand for public web access to NASA's electronic content has made an Internet presence a necessity for NASA. Naturally, a credible, effective web presence that inspires the general public and informs our employees, customers and partners requires a thoughtful hosting decision. NASA is understand/ably concerned with the effectiveness and functionality of our sites, as well as security, privacy and return on investment. Like the first instance of Inside NASA, the first instance of My NASA will be internally hosted.

B. VENDOR MANAGEMENT AND UPGRADE PATH

The fullest participation of the vendor in the development and deployment will lead to an increased, value-added support level. This means that the vendor should be engaged at all critical levels, including product management, technical development and product support. Measurable, bilateral benefits can be achieved from the onset. Future planning will be best achieved through periodic exchange under a mutual non-disclosure agreement (NDA).

Once a decision his made to externally host the NASA Portals, there must be a willingness to work collaboratively and synergistically with our outside vendors. Full participation of the vendor in the development and deployment will increase value-added support. The vendor must be engaged at all critical activities, including product management, technical development and product support. Measurable, bilateral benefits can be achieved from the onset. Future planning will be best achieved through periodic exchange under a mutual non-disclosure agreement (NDA).

The commercial portal environment is not stable but this risk is well known and can be managed through a periodic technology survey. The results need to be propagated to the entire organization. No one method or "industry guru" can be relied upon to give sufficient results in this area. If the current vendor shows signs of "slipping" then this risk needs to be clearly identified and tracked.

Enterprise Information Portals will continue to grow rapidly in functionality. NASA is just beginning to understand what is possible by combining universal communication, business intelligence and workflow management. We must periodically gauge the portal marketplace and assess any potential risk of the current solution with new solutions being offered.

C. VENDOR CAPABILITY DEMONSTRATIONS

Hosting vendor capability demonstrations using a subset of NASA-specific content provides both parties with valuable insight. It should <u>not</u> be done in a way that "exercises" any vendor or misleads them. Feedback should be prompt and fair. This is a reasonable bi-annual activity.

Once a year, a (very) limited development license for the newest portal technology leader should be acquired and rapidly prototyped using applicable content. A succinct recommendation needs to be given to the responsible development organization and funding organization.

D. TECHNICAL TRAINING FOR PORTAL STAFF AT CENTERS

Developers responsible for programming data channels and the general presentation layer will be highly leveraged if they attend the vendor's training classes. This will not obviate the need for an on-site engagement with the vendor's professional services consultant.

Administrators responsible for the portal configuration, channel presentation and sub-portal control will benefit from vendor training classes. Coordinated distributed management will be required, given the size of the job.

The development organization, along with the organizations providing the authentication service, search services and eventually content management services, are responsible for producing and maintaining the Help Desk curriculum.

Table 1. Examples of skills needed for web site versus portal developers

	WEB SITE	PORTAL
HTML	✓	✓
WEB SITE DESIGN (INCLUDING HUMAN FACTORS)	✓	✓
CONTENT MANAGEMENT	✓	✓
SEARCH ENGINES		✓
DATABASE SKILLS (RELATIONAL OR OBJECT-ORIENTED)		✓
PORTAL TECHNOLOGIES (INCLUDING INTELLIGENT AGENTS)		✓
PROJECT MANAGEMENT AND VENDOR RELATIONSHIPS		✓

E. SERVICE LEVEL AGREEMENTS FOR INTERNAL SERVICE AND VENDORS

One of the most important and often overlooked items in working with any outsourced operation is a solid service level agreement (SLA). An SLA is the contract between the managed hosting provider and the company buying the service. It should detail who does what, what is expected and assumed, and what type of services (quantity and quality) are expected to pass between the two parties.

The ***hardware*** and ***software*** issues focus on:

- Separation of the development and operations platforms
- Separation of the data and the application
- Help desk support for technical issues, with a shared knowledge base of problem resolutions
- Security
- Configuration management for software builds and for regression testing purposes
- Regular replenishment of hardware and upgrades to software

1. Administration

Administration of the portal should not a full-time job. However, an administrator needs to be available during times of critical changes in the data channels, sub-portal changes or for emergency changes. An SLA would specify the following

- Administrator accessibility during normal business hours through the four time zones

- Integrated into an institutional problem reporting system, such as Remedy
- Tight integrated into the hosting service

2. Help Desk

The Help Desk SLA should be written to reflect a complete and integrated service provider. SLA items should include (in no particular order)

- Staffing during normal business hours through the four time zones.
- Integrated into the institutional problem reporting system, such as Remedy
- Tight integrated into the hosting service
- Single point of contact for users seeking help

3. Hosting

The Hosting SLA should be written to reflect a complete service provider. SLA items should include (in no particular order)

- Availability (nominal, MTBF, MTTR)
- Performance
- Capacity
- Capacity upon demand
- Security (physical, system and network)
- Metrics reporting
- System administration
- System monitoring
- Backup/restore
- Configuration management
- License management
- Fail-over
- Load balancing
- Operating system upgrades
- Application software upgrades
- Hardware
- Upgrades

F. SYSTEM ANALYSIS AND TESTING

System testing is differentiated from usability testing or design in that this testing ensures that the portal behaves as it is designed, but does not necessarily contribute to the design. Testing needs to be carried out in order to assure performance and proper management of the content. A thorough testing plan needs to be developed based on the applications chosen and the expected uses to which they will be put. Developers need to be involved in: (1) working with a test engineer to develop the types of tests, and (2) setting up the processes by which the test data will be analyzed and acted upon. For example, NASA will need to determine if there are certain expectations for the portal (such as it needs to handle 1,000,000 transactions per hour during peak mission encounters) and when problems occur a specific, named individual is available to help (such as a developer or ASP employee to be paged when there is a hardware or software failure).

Some of the types of testing include web traffic analysis, regression, load and stress, and performance test and monitoring. Some of the testing requirements are noted in the Appendix.

These tools, such as WebTrends, Astra Site Manager, and WebAlizer can be used throughout the development and operational life cycle. They can help the developers determine how the portal is being used: how many times people visit a specific page ("hits"), the length of time spent on each page, what pages or images people are downloading, and how people are navigating through the portal. Such analysis is important so that the portal and channels can be made as efficient and useful as possible. By noting which browsers customers are using, we can optimize the portal for Netscape Navigator, Internet Explorer, or others based upon actual customer usage. These tools also keep track of the search strings customers use so that we can tweak our taxonomies and help guide customers to their product choices as quickly as possible. These tools will be used periodically to check all intranet web sites. Broken links and download times that exceed a specific threshold will be reported to the site developers so that appropriate fixes can be made.

These products can also yield information that is useful for the developers to keep the portal running clean and smooth. Annoying errors, such as "404 Page or File Not Found" or "403 Forbidden Access", broken links, and out-of-date referring pages can be dealt with before the customers see them. Developers can monitor the number of visitors to predict if performance threshold are close. Some of the security monitoring includes the ability to detect unauthorized access attempts and deny or allow visiting search engine spiders.

Once the applications are moved into the integration and testing environment either in-house at Palmers or at the ASP, two tests will be required: a basic regression test and performance testing that will include scripts to verify terminal and disk activity, load imbalances between the processors, the number of operations per second, CPU/disk/memory usage, and system process performance (protocol). Cron jobs will be set up to run these scripts on a regular basis, with the appropriate development staff members being notified if a problem occurs or a threshold is being approached.

G. MAINTENANCE AND SUSTAINING OPERATIONS

There will be regular instances of broken links and abused interface agreements. In order to maintain usability and credibility, this bit rot needs to be fixed on a daily basis. It needs to be a FTE that works well with the evolving content management system. The user has expectations of what will happen as they navigate through NASA’s web space. Understand the behavior of the user—monitoring the number of minutes visitors stay at a site and the number of times they return, cleaning up dead and broken links, and keeping download times to a minimum.

Bugs reported through the Help desk or the integrated problem reporting system need to be evaluated and prioritized for repair, testing and release through the CM system.

Within a development cycle, the development team needs to refresh the presentation capabilities of the most visible data channels. This is not to be confused with the content management system.

Using the channel usage metrics, data channels need to be evaluated for ROI. Heavily used channels need the most attention and deserve investment capital for maintenance and improvement. Underutilized or significantly poor performing channels should be deleted and replaced by ones requested by user survey.

7. NEAR-TERM RECOMMENDATIONS

A. DEPLOYMENT

The portal is planned for a two-phased deployment in order to mitigate technical and programmatic risk and ensure the best possible experience for the users.

The primary aspects that should be **centralized** are COTS integration, licensing, and training. Content management, user training, and focused roll-outs would be **distributed**. Hosting would be on a small number of distributed servers. Integration of the search capabilities is achievable, but a technical challenge.

B. STRATEGIC PLANNING

There are two primary mechanisms for deploying the portal across the Agency: Center by Center and/or community by community. For example, JPL will be implementing a Center-wide approach and all (or most) users will experience a content-rich, populated portal when they begin (think of it as the broad, shallow approach). Building upon this model, communities across Centers or at other Centers could build entrances to project- or community-specific information (think of it as the deep, but focused approach). Eventually, many areas and interests would have a depth of knowledge available.

Once the portal is initially deployed, there is no "webmaster", although there will probably need to be system administrators at the hosting Center and perhaps a Center programmer and administrator to track the logs and deal with new groups. This should be done along the lines of a distributed model, where there is an administrator for each Center. We will identify some basic channels (news, research, and project information) that are needed at the beginning and then establish a group of approved Agency publishers for each channel. The user then has the option of subscribing to news from each of these channels.

Each of the Centers' Public Affairs areas should be a publisher and whatever story appears on their home page should also appear as a headline in the "News from the Centers" channel. Both projects and communities would be useful as they bring together different sets of people that are possibly working on similar problems without an awareness of each other, although we want to be mindful of appointing some kind of moderator or sponsor for each major channel.

Table 2. Sample Measurements for Success

TANGIBLE BENEFITS	METRIC
Higher customer satisfaction	Satisfied user index
Elimination of rework	Number of hours saved
Enhanced knowledge capture	Number of information objects captured in repository
INTANGIBLE BENEFITS	INDICATOR
Improved knowledge sharing	Knowledge organization index
Better understanding of legacy systems	Captured business rules and metadata
Better understanding of business processes	Captured business rules and metadata
Improved business intelligence	Organizational IQ
Improved reputation	Increase in web traffic, user feedback

C. USER COMMUNITIES

The user bases include project and mission teams, communities of practice (any group interested in the same topic), and job groupings (people who do similar work). It would useful to look at the workforce profile for NASA to target specific kinds of research the Agency is presently involved in and that the biggest numbers of employees are engaged in.

Marketing can be done using a number of vehicles, including the portal itself. We have already presented to and started to prepare a network of internal communicators and webmasters across the Centers. The first planned communities and their sponsoring organizations include:

- Standards (Code AE—Weinstein)
- APPL (Code FT—Hoffman)
- e-Learning (Code F—McElwee)

D. PHASE 1

The primary thrust of Phase 1 of the portal will be to develop and deploy an Agency-wide intranet and public portal prototype focusing on the improved access to agency information resources. This will build upon the successful portal pilot developed at JPL using Sun's iPlanet product and Code FE's SpaceLink search capability run by MSFC on Ultraseek software (see Figures 13-15).

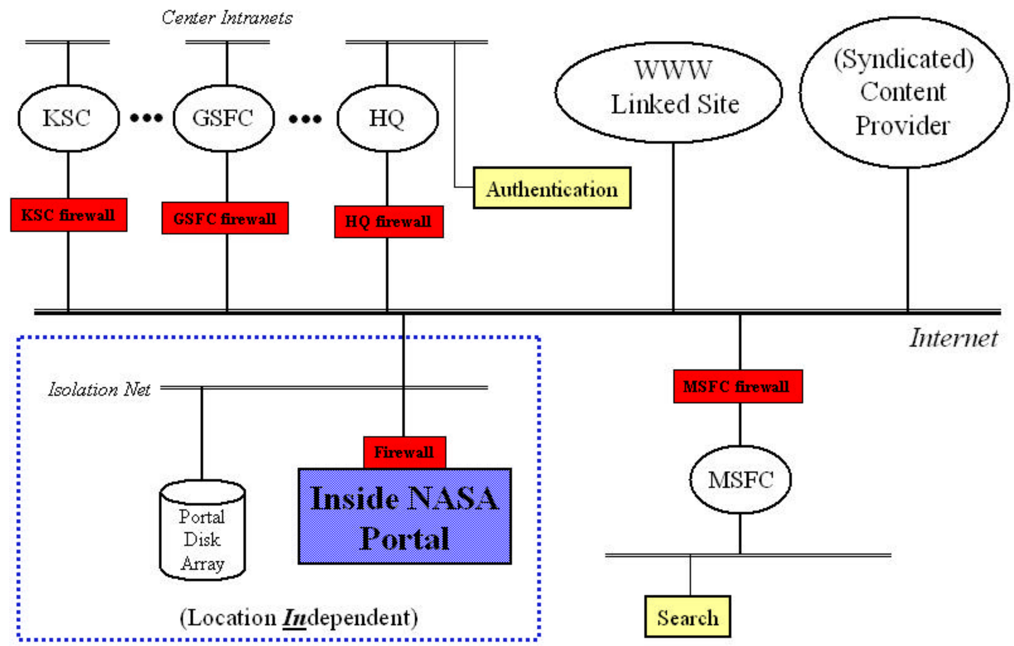

Figure 13. Phase 1: InsideNASA Hardware Architecture

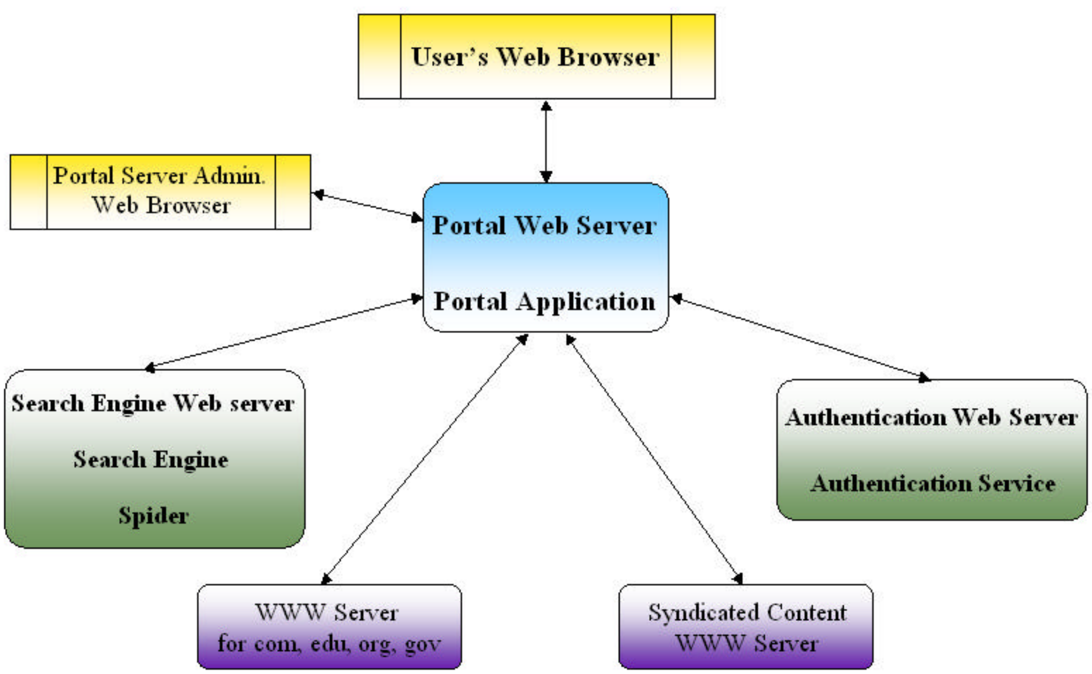

Figure 14. Phase 1: InsideNASA Software Architecture

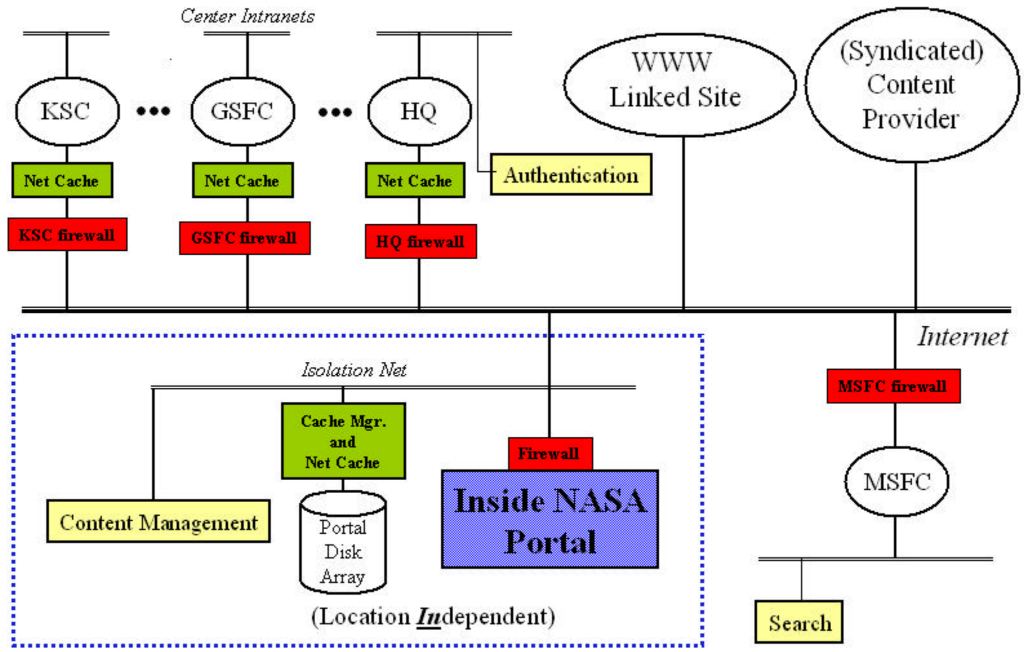

Figure 15. Phase 2: InsideNASA architecture

The Appendix shows a requirements matrix for both the InsideNASA and MyNASA portals, as well as initial prioritization of those requirements for the phased delivery. Requirements will continue to be gathered throughout the initial start up of the project. At the time of the Requirements Review, the requirements for each phase will be baselined.

Authentication will be enabled by a best practice capability already deployed to thousands of NASA employees within several centers and mission areas. The capability provides two-factor authentication using a hardware token and is based on the RSA SecurID product. The authentication service will be provided by the Secure Nomadic Access (SNA) project.

To comply with Agency initiatives in eliminating clear text re-usable passwords, and to provide easy to use, 2810 compliant strong authentication for users accessing the portals, the project intends on leveraging the work already completed under a Headquarters initiative and provide customers with strong authentication to a portal. Strongly authentication means that the user needs two of the following three things (1) Something they have, (2) Something they know, (3) something they are. For this project NASA users that require anything other than public access will be issued a hardware token that will automatically display their passcode every minute. The user when logging in will be challenged for their name and passcode. To enter their passcode they simply append a pre-assigned PIN to whatever passcode is being displayed to gain entry to the portal. This approach will not only provide excellent security and leverage the existing ACE infrastructure across NASA but positions us to make use of *roles based* access to information instead of simple *rules based*. Because of its modular and flexible architecture, this approach will support numerous governance models ranging from Center specific to Agency wide.

We suggest that existing directories be examined for their use in a first-effort portal prototype. The logical place to look for directories that are already serving the public well is the NASA home page and its underlying taxonomies. Study of the NASA homepage yields taxonomies reflecting some of the following topical groupings: mission names, NASA educational themes and work by Center. In addition, the Top Search terms page seems to reflect much searching by planet name.

Taking two or three top-level taxonomies from these sources would represent a first cut look at NASA web content. Although it is severely constrained, it would be a doable goal for a near term prototype effort until more in depth work could be done. The NASA Webmasters have already identified the top Agency sites. It wouldn't be difficult to approach site designers and get their support on a simplified metadata tagging schema as a first step at Agency-wide controlled vocabularies.

A relatively brief set of tags developed from the Dublin Core metadata set can be adopted by Webmasters for universal implementation. This retrofitting effort would be a great improvement over the current state of NASA web space.

Webmasters want their sites to be found. Shrewd ones do all they can to increase their visibility at the top commercial search engines. In the same way we can encourage consistency in metadata implementation by creating a web site registration process for the NASA Search site or expanding the existing one operated by MSFC for Code FE's SpaceLink. Webmasters with new sites that want to be accepted into the NASA Search catalogue would be asked to complete an online registration form that includes metadata fields designed to slot sites into the appropriate directory category.

Due to the fact that implementing a production content management system is neither quick nor cheap, coupled with the fact that there is an effort to roll out a first phase NASA portal in six to twelve months, it is in the best interest of the Agency to invest in some more time in research regarding web-based content management. This research needs to be combined with the final Web services model developed by the Web Management Services (SRR 67) team to develop requirements for a content management system. Further research should continue both in vendor markets and industry trends, as well as, in lessons learned regarding business process associated with the initial phase of the NASA Portal project.

The detailed description of the implementation steps is shown in the schedule noted below.

In order to produce and maintain a NASA portal correctly it is imperative to have both the right management organization, as well as the right technology to manage the vast amounts of content that exists within the Agency.

Table 3. Proposed Implementation Steps for the Phase 1 Portal

NASA PORTAL PHASE 1	
REVIEWS	REQUIREMENTS REVIEW
	SECURITY REVIEW
	CRITICAL DESIGN REVIEW
	OPERATIONAL READINESS REVIEW
	INSIDE NASA PILOT ORR
	MYNASA PILOT ORR
	MISSION PI PILOT ORR
	LESSONS LEARNED REVIEW
COMMUNICATIONS AND ROLLOUT	AGENCY, CENTER, WEBMASTER, AND PUBLISHER COMMUNICATIONS
	TRAINING FOR PORTAL DEVELOPERS, PUBLISHERS, AND USERS
	ONGOING CUSTOMER FEEDBACK
SYSTEM ENGINEERING	BASELINE FUNCTIONAL REQUIREMENTS
	FUNCTIONAL REQUIREMENTS DOCUMENT
	INTERFACE CONTROL DOCUMENTS
	· INTEGRATE NITI RESULTS
	· ESTABLISH INTERFACES TO EXISTING SYSTEMS (PORTAL, SEARCH, SECURE NOMADIC ACCESS, AND COLLABORATIVE TOOLS)

PORTAL MANAGEMENT	PROCURE HARDWARE AND SOFTWARE
	ESTABLISH SERVICE LEVEL AGREEMENT FOR HOSTING AND HELP DESK
	PHYSICAL SITING
	PROGRAMMING AND SECURITY
	MAKE DEVELOPMENT ENVIRONMENT AVAILABLE TO DISTRIBUTED DEVELOPERS
	SYSTEM TESTING, MAINTENANCE, AND REPAIR
INFORMATION ARCHITECTURE	CREATE MAP OF NASA INFORMATION SPACE AND INFORMATION FLOW DIAGRAM
	CREATE TAXONOMY AND IDENTIFY AND INTEGRATE EXISTING TAXONOMIES
	INTEGRATE NASA THESAURUS
	ESTABLISH AND TEST TAXONOMY
ADOPT METADATA STANDARDS	IDENTIFY STANDARDS AND SELECT
	IMPLEMENT IN TWO KEY SYSTEMS
	IDENTIFY OR CREATE METADATA REGISTRY
CONTENT MANAGEMENT	DEVELOP PORTAL MOCK-UPS
	IDENTIFY INTERNAL PUBLISHERS, KEY CONTENT AND SYSTEMS, AND SYNDICATED CONTENT
	CREATE PUBLISHERS' AGREEMENT
	ESTABLISH COMMUNITY CHANNELS (STANDARDS, APPL, E-LEARNING, OTHERS)
INVESTIGATE MANAGED HOSTING SERVICES FOR PHASE 2	IDENTIFY POSSIBLE CANDIDATES, SITE VISITS
	DEFINE MUTUAL REQUIREMENTS, AND MONITOR METRICS AND MANAGE CONTENT
	ESTABLISH CONTRACT IF APPLICABLE
NASA PORTAL PHASE 2	SCALE EXISTING HARDWARE AND SOFTWARE
	MOVE TO MANAGED HOSTING SERVICE (OPTIONAL) (PARALLEL OPERATIONS)
	ENHANCE SEARCH CAPABILITY
	INTEGRATE TO SECURE NOMADIC ACCESS SERVICE
	ESTABLISH CONTENT MANAGEMENT PROCESSES
	SELECT CONTENT MANAGEMENT TOOLS
	MONITOR METRICS AND MANAGE CONTENT

8. BIBLIOGRAPHY

1. Arevalo, Jose, *IFMP Portal Requirements*, October 24, 2001.
2. Firestone, J. M., "Defining the Enterprise Information Portal," *Executive Information Systems*, July 31, 1999.
3. Gartner Group, *Enterprise Portal Trends Emerge Among Confusion*, April 20, 1999.
4. Hunter, Paul, Bruce Milan, and Johnny Medina, *PI Portal*, Goddard Space Flight Center, Greenbelt, October 24, 2001.
5. Knowledge Management Team, *JPL Enterprise Portal System Requirements Document (SRD), Version 1.0*, JPL Document 18855, Jet Propulsion Laboratory, California Institute of Technology, April 5, 2000.
6. Ojakaar, Erik and Jared M. Spool, *Getting Them to What They Want*, User Interface Engineering Reports: The Best Practices Series, October 2001.
7. Pino, Chris, *eNASA Web Portal System Requirements Document (SRD-01)*, Version 1.0, NASA, Washington DC, April 2001.
8. Rasco, Patrick and Justin Jackson, *MSFC Content Management Research and Requirements Definition,* MSFC, Huntsville, Alabama, 2001.
9. Rosenfeld, Louis and Peter Morville, *Information Architecture for the World Wide Web*, O'Reilly, 1998.
10. Spaceport Services, *Kennedy Space Center Portal Overview Information*, KSC, Cape Canaveral, Florida, October 22, 2001.

APPENDIX. FUNCTIONAL REQUIREMENTS LISTING AND PRIORITIES FOR NASA PORTALS FOR INTERNAL AND EXTERNAL AUDIENCES

	Inside		External	
	Phase 1	Phase 2	Phase 1	Phase 2
I. Functional Requirements				
A. Server Side Requirements				
1. Server Platform				
a. Unix server running Sun Solaris	M	M	M	M
b. Unix server running Linux	O	O	O	O
c. Intel-based server running Microsoft Windows NT Server, version 4.0	O	O	O	O
2. Protocol Requirements				
a. Transport Control Protocol / Internet Protocol (TCP/IP)	M	M	M	M
b. Hypertext Transfer Protocol (http), version 1.1	M	M	M	M
c. Secure-Hypertext Transfer Protocol (https)	M	M	M	M
d. Network News Transport Protocol (NNTP)	M	M	M	M
e. Web-based Distributed Authoring and Versioning (WebDAV)	P	P	P	P
f. Lightweight Directory Access Protocol (LDAP)	M	M	M	M
3. Content Source				
a. HTML, version 4.01 and below	M	M	M	M
b. XML	M	M	M	M
c. XHTML	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

d. Plain Text (.txt)	M	M	M	M
e. Application Formats				
1. MS Office 97 (Win), Office 2000 (Win), Office 98 (MAC)	M	M	M	M
2. Word (.doc)	M	M	M	M
3. Excel (.xls, .xlb)	M	M	M	M
4. PowerPoint (.ppt)	M	M	M	M
5. Portable Document Format (.pdf)	M	M	M	M
6. Postscript Format (.ps)	P	P	P	P
7. STEP-Compliant CAD files	P	P	P	P
4. Syndication Content Formats				
a. Information Content and Exchange (ICE), World Wide Web Consortium (W3C) Note 26 October 199	P	M	P	M
5. Databases				
a. Java Data Base Connectivity (JDBC)	O	M	M	M
b. Open Data Base Connectivity (ODBC)	O	M	M	M
6. Portal System Integration and Methods				
a. Tools for integrating with existing NASA and external data sufficient for each defined user group's list of critical COTS, GOTS, and custom applications.	M	M	M	M
b. Open standards API				
1. XML	M	M	M	M
2. C++	M	M	M	M
3. Java	M	M	M	M
4. Perl	M	M	M	M
5. JDBC	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

6. ODBC		M	M	M	M
c. Software Development Kit		M	M	M	M
d. Channel oriented content publishing tools		M	M	M	M
e. Metadata		M	M	M	M
f. Data preparation tools					
1. Web accessible		M	M	M	M
2. Transparent enough for a non-technical user to learn quickly		P	M	P	M
g. Application Support					
1. Organic work group and community support	Policy/directory driven configurable applications which are built in to the portal				
a. Threaded discussions		O	P	O	P
b. Forums		O	P	O	P
c. Bulletin boards		O	P	O	P
d. Calendars		O	P	O	P
e. List serve		O	P	O	P
f. News		O	P	O	P
g. File archives		O	P	O	P
h. Email		O	O	O	O
2. Organic IFM application support		O	O	P	M
3. Bundled, public domain, or commercial application interfaces		O	P	P	M
4. Web application (HTML, JAVA, JAVA Script, XML) support via browser		M	M	M	M
5. Custom Applications	The system shall provide a mechanism to allow administrators to develop custom web-based applications that run within a data channel window	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

B. Security Requirements						
a. Meets NASA NPG 2810 security requirements			M	M	M	M
b. Meets center perimeter and local CCB requirements						
	1. HQ		M	M	P	P
	2. JPL		P	M	M	M
	3. Johnson		P	M	P	M
	4. Kennedy		P	M	P	M
	5. Langley		P	M	P	M
	6. Goddard		P	M	P	M
	7. Marshall		P	M	P	M
	8. Stennis		P	M	P	M
	9. Dryden		P	M	P	M
	10. Glenn		P	M	P	M
C. Client (End-user) Requirements						
1. Client User Access						
	a. Configurable by userID, domain name, IP address, token id				M	M
	c. Internet Protocol Address Access Control	Client access shall be optionally configurable by the administrator based on Internet Protocol (IP) address	M	M		
	d. Domain Name Access Control	Client access shall be optionally configurable by the administrator based on Domain Name	M	M		
	e. Single Sign-On	The system shall be able to provide single sign-on capabilities to any userid/password/token id protected public NASA applications and services	O	M	P	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

f. Token ID	ACE server authentication option	P	M	M	M
2. Client User Interface					
a. Internet Explorer (IE) 4.0 or Netscape 4.0 or higher	The system client user interface must be accessible through the Netscape Communicator or the Microsoft Internet Explorer (IE) version 4.0 (or later) browser or equivalent on MAC and UNIX platforms	M	M	M	M
b. Plug-in requirements	The end-user client interface shall work without the use of non-bundled web browser plug-ins	P	M	P	M
c. Java/Java Script	The end-user client interface may require the use of Java and/or Java Script on the browser	O	O	O	O
d. Windows 95/98/ME Functionality	The end-user client interface shall provide browser based functionality for the Windows 95/98/ME desktop computing platforms	M	M	M	M
e. Windows NT/2000/XP Functionality	The end-user client interface shall provide browser based functionality for the Windows NT/2000/XP desktop computing platforms	P	M	M	M
f. Apple Macintosh Functionality	The end-user client interface shall provide browser based functionality for the Apple Macintosh desktop computing platform	M	M	M	M
g. UNIX Functionality	The end-user client interface shall provide browser-based functionality for the UNIX desktop computing platform	P	M	M	M
h. Palm OS Functionality	Strong protocol and application support	O	P	O	P

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

i. Windows CE Functionality	Strong protocol and application support	O	P	O	P
3. Client Functions					
a. Multiple channels in single page	The system shall allow the display of multiple data channel windows within a single web page	M	M	M	M
b. Browseable directory of NASA web sites	The system shall provide a data channel window, tab, or equivalent user-interface element that accommodates a browseable directory of NASA web sites	M	M	M	M
c. User configurable start page	The system shall provide a personalized start page for each user that is customizable according to the preferences of each user and allows the opportunity to choose the content topics and interactive services to embed into the page	M	M	M	M
d. Log out state management	The system shall keep track of each user's state based on last log out time	M	M	M	M
e. Feedback method	The system shall provide a mechanism to provide feedback and comments	M	M	M	M
1. Link to user's email	The system notification method shall include but not be limited to notification to an external (outside the NASA portal) or internal e-mail account specified by the end-user	P	M		
2. Portal event notifications via email	The system shall allow end-users to receive notifications based on event triggers (a notification is sent when an event occurs, such as when a data channel receives a new object)	P	M		

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

f. Search and directories	The system shall support the current headquarters public search	M	M		
1. Public search engine	The system shall integrate with the current HQ public search engine				
2. Alternate search engines	The system shall provide strong support for integrating secondary engine architectures to address specialized requirements	P	M		
3. Metadata support	Well integrated open standards based including but not limited to Dublin Core	M	M		
4. Directories					
a. Directory hierarchy	The browse directory shall display web sites in an administrator-configurable hierarchy provided by NASA				
b. Directory support	Multiformat, multilevel authored directories with flexible, metadata driven end user and group tools	P	M		
c. Category Tree/Taxonomy	The system shall provide the capability to create a category tree, or taxonomy, which allows end-users to search by category of information rather than by a keyword				
5. Required metadata fields	All content will be cataloged with metadata tags to support search, directory, and channel data assignment and retrieval				
a. Title Information Maintenance	The system shall maintain (although not necessarily simultaneously display) the title information (if available) for each NASA web site or other NASA	M	M		

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

	data objects contained in the browse directory				
b. URL Information Maintenance	The system shall maintain (although not necessarily simultaneously display) the URL information (if available) for each NASA web site or other NASA data objects contained in the browse directory	M	M		
c. Description Information Maintenance	The system shall maintain (although not necessarily simultaneously display) the DESCRIPTION information (if available) for each NASA web site or other NASA data objects contained in the browse directory	M	M		
d. Keyword Information Maintenance	The system shall maintain (although not necessarily simultaneously display) the KEYWORD information (if available) for each NASA web site or other NASA data objects contained in the browse directory	M	M		
e. Categorization Information Maintenance	The system shall maintain (although not necessarily simultaneously display) the CATEGORY information (if available) for each NASA web site or other NASA data objects contained in the browse directory	M	M		
f. Access Restriction Information	The system shall maintain (although not necessarily simultaneously display) the ACCESS RESTRICTION information (if available) for each NASA web site or other NASA data objects contained in the browse directory	M	M		

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

g. Updated Record Information	The system shall maintain (although not necessarily simultaneously display) the date that the last record was updated information (if available) for each NASA web site or other NASA data objects contained in a browse directory	M	M		
h. Layout	Wide range of user or administrator configurable attributes to include:				
1. Data channel layout	Channel layout shall be flexibly specified	M	M	M	M
2. Content channels within channel layout	User configurable attributes of the start page shall at least include the content within specific data channel windows	M	M	M	M
3. Character format	User configurable attributes of the start page shall at least include the font face, size and color	M	M	M	M
1. System-Level Change Notification	The system shall provide a mechanism for end-users to automatically receive notifications of system-level changes as specified by administrators				
2. Data Channel Change Notification	The system shall provide a mechanism for end-users to automatically receive notifications of changes within individual data channels				
3. Data Channel Sub-Elements Change Notification	The system shall provide a mechanism for end-users to automatically receive notifications of changes to sub-elements within individual data channels, such as changes to a document file				

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

4. Notification Receipt Control	End-users shall be able to control which notifications, if any, to which they subscribe				
5. Time Period Notification	The system shall allow end-users to receive notifications based on time period (a summary of changes is sent at a user specified time period)				
k. Security and privilege roles	The system and all its data elements shall be configurable to support all NASA NPG 2810 and all other pertinent policies and policy guidance for the following rolls				
II. Administrative Requirements					
A. Configurable start pages linked to user roles		M	M	M	M
B. Content refresh and management					
1. Publishing Rights and Approval Processes	The system shall support distributed, hierarchical user/group/role assignment of publishing rights	M	M	M	M
2. Role-based Access	All content will be accessed through an integrated role-based privilege assignment system	M	M	M	M
3. New and Updated Content Retrieval	The system shall provide the capability to periodically gather fresh content, and assemble this content for users to view on the portal or via e-mail updates	M	M	M	M
4. Required vs. Optional Data Channels	The system shall accommodate both required data channels that the user cannot remove, and optional data channels that the users can choose to display or not	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

5. Filters	The system shall support the use of filters that will use business rules to determine where content belongs in the portal, to automate the process of adding new content	M	M	M	M
6. Exclusive Content Control	The system shall provide a mechanism for generating reports about logs such as traffic to the system based on both total hits and unique visitors	M	M	M	M
7. Online Content Sources	The system shall allow the user easy access to information feeds including public domain, gratis commercial services, or commercial services	O	O	M	M
8. NASA-Subscribed Sources	The system shall allow the user easy access to syndicated news feeds from NASA-subscribed content sources	O	O	M	M
9. Channel permission structure	The system shall support a flexible, easily administered channel permission structure				
C. Web-based portal management interface		M	M	M	M
D. Flexible scheduling and assignment of labor intensive administrative tasks		P	M	P	M
E. Channel permission structure		M	M	M	M
F. Visual style configuration					
1. Font Appearance	The system shall allow an administrator to control the system-wide appearance of the Font style (e.g., bold, italics, underline, etc,); Font face, Font size, and Font color web page attributes	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

2. Header Content and Appearance	The system shall allow an administrator to control the system-wide appearance of the header content web page attribute	M	M	M	M
3. Footer Content and Appearance	The system shall allow an administrator to control the system-wide appearance of the footer content web page attribute	M	M	M	M
4. Configuration for New Users	The system shall allow an administrator to configure default start pages for new users according to user type or any other administrator-defined group	M	M	M	M
G. Required vs. Optional Data Channels	The system shall accommodate both required data channels that the user cannot remove, and optional data channels that the users can choose to display or not				
H. Frequently Accessed Content Storage	The system shall provide a mechanism to store frequently accessed content and services in memory on the web server in order to maximize scalability and provide users with quick retrieval of the portal page	O	M	P	M
I. User Authentication	The system should support all of NASA's user authentication systems				
1. Kerberos4				M	M
2. Kerberos5				M	M
3. NT domain		M	M	M	M
4. Token ID		M	M	M	M
5. X.509 Protocol	The system should support Public Key Infrastructure (PKI) authentication as specified by the IETF in the X.509 protocol	M	M		

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

J. Users and Groups Administration	The system shall support distributed, hierarchical user/group/role administration. A granular security system be provided to allow administrators to assign different levels of administrative privileges to users throughout the portal, so a distributed community of NASA personnel may manage the portal	M	M	M	M
K. Reporting					
1. Generating Log Reports		M	M	M	M
2. Usage Tracking	The system shall provide a mechanism for tracking usage of specific data channels	M	M	M	M
L. On-Line Help	The system shall supply context-sensitive on-line help to the user	M	M	M	M
M. Web Material Archival	The system shall support a content archival process consistent with all relevant records management policy	M	M	M	M
N. Web Material Retention	The system shall support a content retention process consistent with all relevant records management policy	M	M	M	M
III. Performance Requirements					
A. Multiple User Account Support	The system shall support user account access (having a user name, password) for all public end-users of NASA's web portal.				
1. 10,000 users		M	M	M	M
2. 50,000 users		P	M	O	M
3. Unlimited users (growing server resources as required)		P	P	P	P

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

B. Concurrent User Support	Active users are defined as being currently logged on to the system and using a portal web client				
1. 500 active users		M	M	M	M
2. 5,000 active users		P	M	O	M
3. 50,000 active users		P	P	P	P
C. Concurrent Action Support	These are actions being performed on the server to support users (i.e., log-on, authentication, password change, dynamically generating HTML pages, down load documents or process queries)				
1. 100 active processes		M	M	M	M
2. 500 active processes		P	M	O	M
3. 5000 active processes		P	P	P	P
D. Hours of Operation	The system shall be available and operational 24 hours per day, 7 days a week with a 9x% up-time matrix				
1. NASA business day	8:00 a.m. EST until 9:00 p.m. EST	M	M	M	M
2. 24/7/365 with routine outages	System always available except for routine and emergency outages with 9x% uptime metric	P	M	P	M
3. Non-stop 24/7/365 operation	System always available except for emergency outages with 99.x% uptime metric	O	O	O	O
E. Scalability	The system shall be scalable to multiple servers to allow tuning system performance to accommodate additional users, increasing concurrent users or concurrent actions	M	M	M	M
F. Third-Party Load Balancing Systems	The system shall accommodate third-party hardware and software load balancing systems	M	M	M	M
G. System Performance Reports	The system shall provide system performance reporting capabilities	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

H. Compatible with network monitoring tools TBD		M	M	M	M
IV. Support Requirements					
A. Technical Support					
1. Technical Support	The vendor shall provide technical support to any NASA administrator or system staff personnel with a NASA domain e-mail address	M	M	M	M
2. Technical Support Hours	The vendor shall provide telephone technical support Monday through Friday, 8:00 am through 6:00 pm Eastern Standard Time (EST)	M	M	M	M
3. Technical Support Responses	The vendor shall respond within two working hours to a telephone technical support call. If the call is received after 2:45 pm EST, then the call may be returned as late as 8:45 am EST on the next NASA working day. It is the preference of NASA to have telephone technical support calls answered by a live voice in an expedient manner	M	M	M	M
4. E-mail Requests for Technical Support	The vendor shall be capable of receiving e-mail requests for technical support	M	M	M	M
5. Acknowledge of Receipt Via Return E-Mail	The vendor's technical support system shall promptly acknowledge the receipt of an e-mail request for technical support via return e-mail to the sender	M	M	M	M
6. On-Site Support	The vendor shall be capable of providing on-site expert technical support	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

7. Fulfilling On-Site Support Request	The vendor shall be capable of fulfilling a request for expert on-site technical support within three (3) NASA working days	M	M	M	M
B. Technical Support Request Tracking System	The vendor shall provide a web-accessible tracking program for all technical support requests	M	M	M	M
C. System Shut Down	The date and time of any required system shut down shall be configurable by authorized administrators at a specified date and time	M	M	M	M
D. Context-Sensitive On-line Help		M	M	M	M
E. Online tutorial and other built-in training materials		P	M	P	M
V. Software Requirements					
A. Host Location	The system software may be hosted at NASA, off-site or at a combination of both, pursuant to compliance with NASA information security requirements (NPG 2810.1)	M	M	M	M
B. Modular Software Architecture	The software shall be able to run unbundled from proprietary software in such a way that different components (I.e., search engine, content management systems) can be easily changed out to allow for other vendors' products to run in tandem with the portal software	M	M	M	M
C. IP Logging	The system software shall log all access attempts to the system by IP address	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

D. Security of information					
1. Security of BRT information	Pursuant to NASA Procedures and Guidelines, "Security of Information Technology", NPG 2810, the consumer portal software will handle BRT Information	O	O	O	P
2. Security of SER Information	Pursuant to NASA Procedures and Guidelines, "Security of Information Technology", NPG 2810, the consumer portal software will handle Scientific, Engineering, and Research (SER) Information	O	P	M	M
3. Security of ADM Information	Pursuant to NASA Procedures and Guidelines, "Security of Information Technology", NPG 2810, the consumer portal software will handle Administrative (ADM) Information	O	P	M	M
4. Security of PUB Information	Pursuant to NASA Procedures and Guidelines, "Security of Information Technology", NPG 2810, the consumer portal software will handle Public Access (PUB) Information	M	M	M	M
VI. Interface Requirements					
A. Open Protocol Support	The vendor shall provide a list of open protocols that are supported and a statement as to what degree they are supported	M	M	M	M
B. EudoraPro Support	The system shall be capable of launching the NASA EudoraPro or other popular e-mail software within the consumer portal	P	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable

VII. Training Requirements					
A. Web Portal Training for NASA Personnel	The vendor shall provide appropriate training to in the design, implementation, administration and maintenance of the portal server and accompanying application to NASA technical personnel and its contractors	M	M	M	M
B. Web Portal Training for Information Publishers	The vendor shall provide appropriate training to information publishers creating data streams for display in portal data channels	M	M	M	M
VIII. Test Requirements					
A. Internal Acceptance Testing	The vendor shall perform internal acceptance testing on the software prior to each delivery	M	M	M	M
B. User Interface Testing	The vendor shall perform user interface testing on the software prior to each delivery	M	M	M	M
C. Performance Testing	The vendor shall perform performance testing on the software prior to each delivery	M	M	M	M
D. Load Level Testing	The vendor shall perform load testing prior to each delivery. The load levels tested shall reflect the load levels listed in section III. Performance	M	M	M	M
E. Load Level Testing Duration	The period for each load level tested shall be a continuous 24 hours	M	M	M	M

M = Mandatory; **P** = Preferred; **O** = Optional; **blank** = No Opinion; **N/A** = Not applicable